

\*\*\*\*\*  
 RELEASE  
 \*\*\*\*\*  
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_n n a : n.a. database search, using Smith-Waterman algorithm  
 Run on: Mon Mar 15 10:55:31 1999. MasPar time 7823.72 Seconds  
 1545 001 Million cell updates/sec  
 Tabular output not generated.

File: >US-09-020-716-5  
 Description: (1:5115) from US09020716 seq  
 Perfect Score: 5115  
 N.A. Sequence: 1 GTTGGAGCTCTCCGATATG TCAAGCTATGCTACACGC 5115  
 Comp: CAACCTCGAGATGATATAG AGTTCGATACCTAGGTTCCG

Scoring table: TABLE default

Gap 5

Mismatch STD : Phase 0: Query 0

Searched: 602357 seqs, 1181590623 bases x 2

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: emb157

1:em\_ba 2:em\_fun 3:em\_htg 4:em\_hum1 5:em\_hum2 6:em\_in  
 7:em\_om 8:em\_ov 9:em\_ov 10:em\_pat 11:em\_ph 12:em\_pl  
 13:em\_ro 14:em\_un 15:em\_vi

Database: genbank110

16:gb\_bal 17:gb\_ba2 18:gb\_htg 19:gb\_in 20:gb\_om 21:gb\_ov  
 22:gb\_pat 23:gb\_ph 24:gb\_pl 25:gb\_pl2 26:gb\_pr1  
 27:gb\_pr2 28:gb\_pr3 29:gb\_ro 30:gb\_st 31:gb\_sts 32:gb\_sy  
 33:gb\_un 34:gb\_vi

Statistics: Mean 12.338 Variance 5.557 scale 2.220

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2594	50.8	3000	32	CVGEM72FP Cloning vector pGEM-72	0.00e+00
2	2596	50.8	3003	32	CVGEM52FP Cloning vector pGEM-52	0.00e+00
3	2596	50.8	4486	32	EVOR116N Expression vector pCOR	0.00e+00
4	2598	50.2	3033	32	CVGEM7LICF Ligation-independent c	0.00e+00
5	2598	50.2	3033	32	CVGEM7LICF Ligation-independent c	0.00e+00
6	2598	50.2	7823	32	AF041436 Cloning vector pVLH-1	0.00e+00
7	2268	44.3	3000	32	CVGEM72FM Cloning vector pGEM-72	0.00e+00
8	2268	44.3	3003	32	CVGEM52FM Cloning vector pGEM-52	0.00e+00
9	2200	43.0	3181	32	CVGEM132P Cloning vector pGEM-13	0.00e+00
10	2200	43.0	3199	32	CVGEM32ZP Cloning vector pGEM-32	0.00e+00
11	2200	43.0	3223	32	CVGEM112M Cloning vector pGEM-11	0.00e+00
12	2200	43.0	3223	32	CVGEM112M Cloning vector pGEM-11	0.00e+00
13	2200	43.0	3223	32	CVGEM112P Cloning vector pGEM-11	0.00e+00

C 14	2200	43.0	4933	32	CVGEM132C	Cloning vector pGEM-13	0.00e+00
C 15	2176	42.5	2743	32	CVGEM32Z	Cloning vector pGEM-32	0.00e+00
C 16	2176	42.5	4283	22	163120	Sequence 1 from patent	0.00e+00
C 17	2176	42.5	4283	22	184496	Sequence 1 from patent	0.00e+00
C 18	2153	42.1	6824	22	A25929	Yeast integration vect	0.00e+00
C 19	2148	42.0	6160	16	EC17DN3	E.coli pT7HspI DNA	0.00e+00
C 20	2148	42.0	9953	16	EC17DN2	E.coli pT7HspI DNA	0.00e+00
C 21	2148	42.0	11371	16	EC17DN1	E.coli pT7HspI DNA	0.00e+00
C 22	2142	41.9	2686	32	XXU17492	Cloning vector pUC-Sce	0.00e+00
C 23	2142	41.9	2746	32	CVGEM42Z	Cloning vector pGEM-42	0.00e+00
C 24	2142	41.9	2964	32	SYNRLSKMV	BlueScribe SK Minus c1	0.00e+00
C 25	2142	41.9	2964	32	SYNRLSKMV	BlueScribe SK Minus c1	0.00e+00
C 26	2142	41.9	3811	32	SYNPS19V	PSPI9 cloning vector	0.00e+00
C 27	2144	41.9	3394	32	AF083409	Cloning vector p4S-Lp	0.00e+00
C 28	2144	41.9	3394	32	AF083409	Cloning vector p4S-Lp	0.00e+00
C 29	2142	41.9	3563	32	XXU04610	Cloning vector p4S-Tp	0.00e+00
C 30	2144	41.9	3604	32	AF083407	Cloning vector p4S-Cm	0.00e+00
C 31	2144	41.9	3604	32	AF083407	Cloning vector p4S-Cm	0.00e+00
C 32	2144	41.9	3617	32	AF083408	Cloning vector p4S-Gm	0.00e+00
C 33	2144	41.9	3714	32	AF083408	Cloning vector p4S-Gm	0.00e+00
C 34	2144	41.9	4236	32	AF062080	Cloning vector p4S-Km	0.00e+00
C 35	2144	41.9	4354	32	AF062082	Cloning vector p4S-Tc	0.00e+00
C 36	2142	41.9	4312	32	AF025747	Cloning vector pUC-Hy	0.00e+00
C 37	2144	41.9	4503	32	AF025747	Cloning vector p4S-Sm	0.00e+00
C 38	2142	41.9	5154	32	XXU17502	Cloning vector p4S-Lp	0.00e+00
C 39	2142	41.9	5153	32	XXU34320	E.coli/K12/yeast/omycetes 1	0.00e+00
C 40	2142	41.9	5153	32	XXU34319	E.coli/K12/yeast/omycetes 1	0.00e+00
C 41	2142	41.9	5646	32	AF025746	Cloning vector pUC-GM	0.00e+00
C 42	2142	41.9	5932	32	XXU34321	Cloning vector p4S-Lp	0.00e+00
C 43	2142	41.9	6395	32	AF025746	Cloning vector pUC-Hy	0.00e+00
C 44	2142	41.9	6894	32	XXU34321	Cloning vector p4S-Lp	0.00e+00
C 45	2142	41.9	7272	32	XXU34321	E.coli/K12/yeast/omycetes 1	0.00e+00

#### ALIGNMENTS

RESULT	1	CVGEM72FP	3000 bp	DNA	SYN	12-FEB-1996
LOCUS		Cloning vector pGEM-72f(+)				
DEFINITION		X65310				
ACCESSION		X65310				
NID		958180				
KEYWORDS		Beta-lactamase, bla gene, cloning vector, lacZ gene; Multiple cloning site; phage II region; promoter.				
SOURCE		unidentified cloning vector				
ORGANISM		unidentified cloning vector				
REFERENCE		artificial sequence; cloning vectors.				
AUTHORS		1 (bases 1 to 3000)				
TITLE		Technical Services				
JOURNAL		Direct Submission				
REMARK		Submitted (23-MAR-1992) Technical Services, Promega Corporation.				
REFERENCE		2 (bases 1 to 3000)				
AUTHORS		2800 Woods Hollow Road, Madison, WI 53711-5399, USA				
TITLE		revised by [2]				
JOURNAL		Klock, C.				
COMMENT		Direct Submission				
		Submitted (28-MAY-1993) Technical Services, Promega Corporation,				
		2800 Woods Hollow Road, Madison, WI 53711-5399, USA				
		This vector can be obtained from Promega Corporation, Madison, WI.				
		Call one of the following numbers for order or technical information:				
		Order or technical 800-356-9526				
		In Wisconsin 800-356-9526				
		Outside U.S. 608-274-4330				
		See <X65300-X65335> for related vector sequences.				
FEATURES		Location/Qualifiers				
source		1..3000				
		/organism="Cloning vector"				
		/plasmid="Plasmid pGEM"				
misc_feature		1				
		/note="7 transcription initiation site"				
misc_feature		10..110				
		/note="Multiple cloning sites"				
promoter		118..140				









```

Cp 2955 GAAAAATAACAAATAGGGGTTCCGGGCACATTTCCCGGAAAAAGTGCACCTCTATCGG 2896
Db 2321 TGTGAATACCCACACAGATCGGTAGGAGAAATACCGCATCCAGCGGAAATTTGTAACGT 2380
Cp 2895 TGTGAATACCCGATACAGATCGGTAGGAGAAATACCGCATCCAGCGGAAATTTGTAACGT 2836
Db 2381 TAATATTTTGTAAATTCGGGTAAATATTGTTAAATCAGGTCAATTTTTTAACCAATA 2440
Cp 2835 TAATATTTTGTAAATTCGGGTAAATATTGTTAAATCAGGTCAATTTTTTAACCAATA 2776
Db 2441 GCGCAAAATCGCAAAATCCCTTAAATCAAAAGATAGACCCAGATAGGTTGACTGT 2500
Cp 2775 GCGCAAAATCGCAAAATCCCTTAAATCAAAAGATAGACCCAGATAGGTTGACTGT 2716
Db 2501 TGTTCAGTTTGGAAACAAGATCCCACTATTAAAGAGGTGACTCCCAACGTCAAAGGGCG 2560
Cp 2715 TGTTCAGTTTGGAAACAAGATCCCACTATTAAAGAGGTGACTCCCAACGTCAAAGGGCG 2656
Db 2561 AAAAGCGGTCTATAGGCGGATGGGCGACTACGTGAACCATCATCCCAAAATCAAGTTTTT 2620
Cp 2655 AAAAGCGGTCTATAGGCGGATGGGCGACTACGTGAACCATCATCCCAAAATCAAGTTTTT 2596
Db 2621 GCGGTGAGGTGCGGTAAAGTCTAAATCGGAACCTTAAAGAGGTGACTCCCAACGTCAAAGGGCG 2680
Cp 2595 GCGGTGAGGTGCGGTAAAGTCTAAATCGGAACCTTAAAGAGGTGACTCCCAACGTCAAAGGGCG 2536
Db 2681 TTGACGGGGAAGCGG 2696
Cp 2535 TTGACGGGGAAGCGG 2520

RESULT ?
LOCUS EVCOR16N 4486 bp DNA SYN 24-MAR-1994
DEFINITION Expression vector pCOR16N (modified from pCOR16 in [3]).
ACCESSION Z29589
VERSION 1
KEYWORDS ampicillin resistance; beta-lactamase; ColE1 origin; expression
vector; nos terminator; phage fl region; rice actin1 promoter.
SOURCE expression vectors.
ORGANISM artificial sequence; expression vectors.
REFERENCE 1 (bases 1 to 4486)
AUTHORS Liu, L., Dasgupta, I., Davies, J. and Hull, R.
TITLE Modified expression vectors for monocot transformation toward virus
resistance
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4486)
AUTHORS McElroy, D., Blowers, A. D., Jones, R. and Wu, P.
TITLE Construction of expression vectors based on the rice actin 1 (Act1)
5' region for use in monocot transformation
Mol. Gen. Genet. 231 (1), 150-160 (1991)
REMARK 92692956
REFERENCE 3 (bases 49 to 304)
AUTHORS Depicker, A., Stachel, S., Dhaese, P., Zambryski, P. and Goodman, H.M.
TITLE Napaline synthase: transcript and DNA sequence
J. Gen. Appl. Microbiol. 1, 561-573 (1982)
REFERENCE 4 (bases 1 to 4486)
AUTHORS Liu, L.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-1994) to the EMBL/GenBank/DBJ databases. Liu, L.,
John Innes Institute, Virus Research, Colney Lane, Norwich, United
Kingdom, NR4 7UH
FEATURES
source
Location/Qualifiers
1..4486
/organism="Expression vector"
misc_feature
1..52
/organism="multiple cloning sites"
misc_feature
3..5
/organism="start codon for gene expression"
terminator
49..304
/organism="nos terminator"
/citation=[3]

```

```

misc_feature 511..967
CDS
/organism="phage fl region"
1153..2013
/organism="beta-lactamase coding region"
/codon_start=1
/transl_table=11
/db_xref="PIR:q452349"
/translation="MSIOHFRVALIPFFAAGCPVFAHPETLVKVKDAEDQLGARVSV
IEDLSNGKILSEFPFPMSTFKVLLCHTLLSPFDAGDGLGPARYSNDLVE
YSVTEKHLLDSMTVELCSAALTMSDNTAANLLTIIGGKELTAFIRHNKSDRTLR
DRPELNEAIPNDERDITLRLKLTIGELLASQQQLDKWEAKVAAPL
LRALPAGFWFIADKSGAGERSGFIIALGPDGKPSPIVVIYITGDAIMDERNSQA
EIGASLIKHW"
misc_feature 2014..2954
promoter
/organism="ColE1 origin"
3256..4486
/organism="rice actin 1 5' region, (act1 promoter)"
/citation=[2]
BASE COUNT 1143 a 1118 c 1048 g 1177 t
ORIGIN
Query Match 50.8%; Score 2596; DB 32; Length 4486;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 639 CGGCTTTCCCGCTCAAGCTCTAAATCGGGGCTCCCTTTAGGGTTCCGATTTAGAGCTTT 698
Qy 2520 CGGCTTTCCCGCTCAAGCTCTAAATCGGGGCTCCCTTTAGGGTTCCGATTTAGAGCTTT 2579
Db 699 ACGGACCTCCAGCGCAAAAACCTTGATTTGGGTGATGGTTCAGTAGTGGGGCATCCGC 758
Qy 2580 ACGGACCTCCAGCGCAAAAACCTTGATTTGGGTGATGGTTCAGTAGTGGGGCATCCGC 2639
Db 759 CTGATAGACGGTTTTCGCGCTTTGAGCTTTGAGCTTCAAGTCTTAAATAGTGGAGCTTT 818
Qy 2640 CTGATAGACGGTTTTCGCGCTTTGAGCTTTGAGCTTCAAGTCTTAAATAGTGGAGCTTT 2699
Db 819 GTTCCAAACTGGAAACACACTCAACCTTATCTGGGTCTATCTTCTGATTTATAAGGGAT 878
Qy 2700 GTTCCAAACTGGAAACACACTCAACCTTATCTGGGTCTATCTTCTGATTTATAAGGGAT 2759
Db 879 TTTCGGGATTTGGGCTTATGGTTAAAAAATGAGCTGATTTAAACAAATTTAAACCGAA 938
Qy 2760 TTTCGGGATTTGGGCTTATGGTTAAAAAATGAGCTGATTTAAACAAATTTAAACCGAA 2819
Db 939 TTTTAAACAAATATTAGCTTTTACAAATTTGCGGTGATGGGTATTTCTCTTACGCTC 998
Qy 2820 TTTTAAACAAATATTAGCTTTTACAAATTTGCGGTGATGGGTATTTCTCTTACGCTC 2879
Db 999 TGTGGGTATTTCACACCGCATACAGGTGGCACTTTTCGGGGAATATGTGGGGAACCGCC 1058
Qy 2880 TGTGGGTATTTCACACCGCATACAGGTGGCACTTTTCGGGGAATATGTGGGGAACCGCC 2939
Db 1059 TATTTGTTTATTTTCTAAATATATCAATATGATGATGATGATGATGATGATGATGATG 1118
Qy 2940 TATTTGTTTATTTTCTAAATATATCAATATGATGATGATGATGATGATGATGATGATG 2999
Db 1119 ATAAATGCTTCAATATATGAAAAAGAGAGATGATGATGATGATGATGATGATGATG 1178
Qy 3000 ATAAATGCTTCAATATATGAAAAAGAGAGATGATGATGATGATGATGATGATGATG 3059
Db 1179 CCTTATTCCTTTTTCGGGCACTTTTGGCTTCTGTTTTCCTTACCGGCAACCGTGGT 1238
Qy 3060 CCTTATTCCTTTTTCGGGCACTTTTGGCTTCTGTTTTCCTTACCGGCAACCGTGGT 3119
Db 1239 GAACTTAAAGATGCTGAAGATCAGTTGGGTGACGAGTGGGTGATGATGATGATGATGAT 1298
Qy 3120 GAACTTAAAGATGCTGAAGATCAGTTGGGTGACGAGTGGGTGATGATGATGATGATG 3179
Db 1299 CAACACGGGTAAAGATCCTTTCAGAGTTTTTCGCCCGGAGAAACGTTTTCCTATGATGAC 1358
Qy 3180 CAACACGGGTAAAGATCCTTTCAGAGTTTTTCGCCCGGAGAAACGTTTTCCTATGATGAC 2399

```



2 (bases 1 to 3033)  
 Haun, R.S.  
 Direct Submission  
 Submitted (19-APR-1995) Randy S. Haun, NIH, NHLBI, Laboratory of  
 Cellular Metabolism, Bldg. 10, Room 5N307, Bethesda, MD 20892-1434,  
 USA  
 Cloning-independent cloning vector. Preparation of the vector for  
 cloning includes linearization with Nari, gel purification of the  
 linearized vector, and treatment with T4 DNA polymerase in the  
 presence of dATP. The target sequence can be amplified using  
 sequence specific primers modified at the 5' end to contain an  
 additional 13 nucleotides complementary to the vector. The forward  
 primer should contain 5'-CTGCTCCGGGCA-3' followed by 12-15  
 nucleotide target-specific sequence. The reverse primer should  
 contain 5'-CTGCTCCGGGCA-3' followed by 12-15 nucleotide  
 target-specific sequence. The amplified sequence should be gel  
 purified and treated with T4 DNA polymerase in the presence of  
 dATP. Annealing of the vector and the amplification product forms a  
 duplex that can be used directly for transformation. Sequences  
 amplified using these primers are also compatible with the  
 pUC19 vector (ATCC 87047) and pGEM-7zf(+) (ATCC 87048).  
 Only in the orientation of complementary ends generated at the  
 cloning site. This vector has been deposited in the ATCC repository  
 (ATCC 87049).  
 Location/Qualifiers  
 1..3033  
 /organism="cloning vector pGEM-7zf(+)/LIC-R"  
 /note="pGEM-7zf(+)/LIC-R. ATCC 87049"  
 /db\_xref="taxon:39794"  
 55..81  
 /evidence="experimental"  
 /product="ligation-independent cloning region"  
 65..70  
 /evidence="experimental"  
 /product="Nari restriction site"  
 complement(151..173)  
 /note="bacteriophage SP6 promoter"  
 complement(193..195)  
 /gene="lacZ"  
 complement(193..195)  
 /gene="lacZ"  
 /note="initiation codon"  
 /codon\_start=1  
 /translation="M"  
 complement(212..212)  
 /codon\_start=1  
 /translation="L"  
 /product="bo3-lacZ gene"  
 /db\_xref="PID:q805880"  
 /translation="MSIQHFRVALIPFFFAAFLPFAHPELVKKDAEDLGARVGY  
 IELDNLGKILSFREPRPMMSTFKVLGGVLSRIDAQGEOLGRHYSONDLVE  
 YSPVTEKHLDGWTVELCSAATMSDNTAANLLITTCGRKELTAFVHMGDHTREL  
 DRPELNEAIPIRNDRTTPVAMATTLKRLTIGELTLASRQQLIDWEDKRVAGPL  
 LRSLAPAGVFALESQNGEPGSGIIAALGDPKPSRVIVYITGSGQAIMDNRQIA  
 EIGASIKHW"  
 complement(2398..2853)  
 /standard\_name="fl origin"  
 join(3017..3033,1..6)  
 /note="bacteriophage T7 promoter"  
 779 a 748 c 769 g 737 t  
 Query Match 50.2% Score 2568; DB 32; Length 3033;  
 Best Local Similarity 96.5% Prod. No. 00e-00;  
 Matches 2592; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
 Db 131 GGGTGTATCATAGCTTGTAGTATCTATATAGTACCTAAATAGCTTGGCTTAATCATG 190  
 Cp 5115 GAGTGTATCATAGCTTGTAGTATCTATATAGTACCTAAATAGCTTGGCTTAATCATG 5056

Db 191 GTCATAGCTGTTTCCTGTTGTAATTTGTTATGCGCTCAATATGCAATCAATCAATCAATCAAT 250  
 Cp 5055 GTCATAGCTGTTTCCTGTTGTAATTTGTTATGCGCTCAATATGCAATCAATCAATCAATCAAT 4995  
 Db 251 CGAAGCATAAAGTGTAAAGGCTTGGGCTGCTTAATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 310  
 Cp 4995 CGAAGCATAAAGTGTAAAGGCTTGGGCTGCTTAATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 4935  
 Db 311 GTTGGCTCACTGCGGCTTTTCAGTCTGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 370  
 Cp 4935 GTTGGCTCACTGCGGCTTTTCAGTCTGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4875  
 Db 371 CGGCAACGCTGCGGCTGCGGCTTTTCAGTCTGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 430  
 Cp 4875 CGGCAACGCTGCGGCTGCGGCTTTTCAGTCTGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 4815  
 Db 431 TGACTGCT 490  
 Cp 4815 TGACTGCT 4755  
 Db 491 AATACGGTTATCCACAGATCAAGGCGGATCAAGGCGGATCAAGGCGGATCAAGGCGGATCAAGGCGG 550  
 Cp 4755 AATACGGTTATCCACAGATCAAGGCGGATCAAGGCGGATCAAGGCGGATCAAGGCGGATCAAGGCGG 4695  
 Db 551 GCAAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 610  
 Cp 4695 GCAAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4635  
 Db 611 CCGTACGAGCATCAACAAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 670  
 Cp 4635 CCGTACGAGCATCAACAAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4575  
 Db 671 ATAAAGTACGAGCGGCTTTCCGCGCTGGAAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 730  
 Cp 4575 ATAAAGTACGAGCGGCTTTCCGCGCTGGAAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 4515  
 Db 731 GCGGCTTACGCGGATACCTGTCGCGGCTTTTCGCGCTTGGGGAAGCGGCGGCGGCGGCGGCGGCGG 790  
 Cp 4515 GCGGCTTACGCGGATACCTGTCGCGGCTTTTCGCGCTTGGGGAAGCGGCGGCGGCGGCGGCGGCGG 4455  
 Db 791 CTCACGCTGTAGGTATCTAGTCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 850  
 Cp 4455 CTCACGCTGTAGGTATCTAGTCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 4395  
 Db 851 CGAACCGCGGCTTACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 910  
 Cp 4395 CGAACCGCGGCTTACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4335  
 Db 911 CCGGCTTAAAGTACAGCTTATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 970  
 Cp 4335 CCGGCTTAAAGTACAGCTTATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4275  
 Db 971 GAGTATGTAGGCGGCTGCTAGCAGTCTGCTAGGCTGCTAGGCTGCTAGGCTGCTAGGCTGCTAGGCTGCT 1030  
 Cp 4275 GAGTATGTAGGCGGCTGCTAGCAGTCTGCTAGGCTGCTAGGCTGCTAGGCTGCTAGGCTGCTAGGCTGCT 4215  
 Db 1031 AAGGACAGTATTTGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1090  
 Cp 4215 AAGGACAGTATTTGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4155  
 Db 1091 TAGCTTGTAGTCCGCGCAACAAACCGGCTGCTAGGCTGCTAGGCTGCTAGGCTGCTAGGCTGCTAGGCTGCT 1150  
 Cp 4155 TAGCTTGTAGTCCGCGCAACAAACCGGCTGCTAGGCTGCTAGGCTGCTAGGCTGCTAGGCTGCTAGGCTGCT 4095  
 Db 1151 GCAGATTACGCGCGAGAAAAAGGATCTCAAGAAAGGATCTCAAGAAAGGATCTCAAGAAAGGATCTCAAG 1210  
 Cp 4095 GCAGATTACGCGCGAGAAAAAGGATCTCAAGAAAGGATCTCAAGAAAGGATCTCAAGAAAGGATCTCAAG 4035  
 Db 1211 TGAGGCTTGTAGTCCGCGCAACAAACCGGCTGCTAGGCTGCTAGGCTGCTAGGCTGCTAGGCTGCTAGGCTGCT 1270  
 Cp 4035 TGAGGCTTGTAGTCCGCGCAACAAACCGGCTGCTAGGCTGCTAGGCTGCTAGGCTGCTAGGCTGCTAGGCTGCT 3975  
 Db 1271 GATCTTCACTAGTCTGCTTAAATTAAGGCTGCTTAAATTAAGGCTGCTTAAATTAAGGCTGCTTAAATTAAG 1330



















```

Db 1781 GTTGGGCGGAGTGTATACATCATATGTTATGGCAGACATGACATAATCTCTTACTGTGAT 1840
Cp 3435 GTTGGGCGGAGTGTATACATCATATGTTATGTTATGTTATGTTATGTTATGTTATGTTAT 3376
Db 1841 GCGATCGGTAGATGCTTTTCTGTGACTGTGTGAGTACTCAACCAAGTCAATCTTGAGATA 1900
Cp 3375 GCGATCGGTAGATGCTTTTCTGTGACTGTGTGAGTACTCAACCAAGTCAATCTTGAGATA 3316
Db 1901 CCGGCGGCGGCGGACCGAGTGTCTTCTCGCGCGGTCAATACGGGATAAATAGTGTATGACA 1960
Cp 3315 CCGGCGGCGGCGGACCGAGTGTCTTCTCGCGCGGTCAATACGGGATAAATAGTGTATGACA 3256
Db 1961 TAGCAGAACTTTAAAGTGTCTCATCATTTGAAACAGTCTTCGCGCGGAAAACTCTCAAG 2020
Cp 3255 TAGCAGAACTTTAAAGTGTCTCATCATTTGAAACAGTCTTCGCGCGGAAAACTCTCAAG 3196
Db 2021 GATCTTACCGCTGTGTGAGATCCAGTTCGATGTAAACCCACTGTGCACCCCACTGATCTTC 2080
Cp 3195 GATCTTACCGCTGTGTGAGATCCAGTTCGATGTAAACCCACTGTGCACCCCACTGATCTTC 3136
Db 2081 AGCATCTTTTACTTTCACCGAGTCTTCTCGCGGTGAGCAAAACAGGAAGGCAAAATGCCGC 2140
Cp 3135 AGCATCTTTTACTTTCACCGAGTCTTCTCGCGGTGAGCAAAACAGGAAGGCAAAATGCCGC 3076
Db 2141 AAAAAGGGAATAAGGCGACACGGAAATGTTGAATACTACTACTCTTCTCTTTTCAATA 2200
Cp 3075 AAAAAGGGAATAAGGCGACACGGAAATGTTGAATACTACTACTCTTCTCTTTTCAATA 3016
Db 2201 TTATGAAGCATTTATCAGGTTATTGCTCTCATGAGCGGATACATATTTGAATGATTTA 2260
Cp 3015 TTATGAAGCATTTATCAGGTTATTGCTCTCATGAGCGGATACATATTTGAATGATTTA 2956
Db 2261 GAAATAAATAAATAGGTTTCTCGCGGCACATTTCCCGGAAAGTCCCACTGTATGGG 2320
Cp 2955 GAAATAAATAAATAGGTTTCTCGCGGCACATTTCCCGGAAAGTCCCACTGTATGGG 2896
Db 2321 TGTGAATACGCGACAGATCGTAAAGAGAAATAACCGATCAGCGA 2368
Cp 2895 TGTGAATACGCGACAGATCGTAAAGAGAAATAACCGATCAGCGA 2848

RESULT 9 CVGEM132P 3181 bp DNA SYN 12-FEB-1996
LOCUS Cloning vector pGEM-132f(+).
DEFINITION Cloning vector pGEM-132f(+).
ACCESSION X65315
NID g58170
KEYWORDS beta-lactamase; bla gene; cloning vector; lacZ gene; multiple cloning site; phage fl region; promoter; pUC/M13 primer binding site.

SOURCE
ORGANISM unidentified cloning vector.
REFERENCE 1 (bases 1 to 3181)
AUTHORS Technical Services.
JOURNAL Direct Submission
REMARK Submitted (23-MAR-1992); Technical Services, Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
AUTHORS revised by (2)
TITLE 2 (bases 1 to 3181)
JOURNAL Direct Submission
REMARK Submitted (28-MAY-1993); Technical Services, Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
COMMENT This vector can be obtained from Promega Corporation, Madison, WI. Call one of the following numbers for order or technical information:
Order or Technical 800-356-9526
In Wisconsin 800-356-9526
Outside U.S. 608-274-4330
See <X65300-X65335> for related vector sequences.
FEATURES
Location/Qualifiers
Source 1..3181
/organism="Cloning vector"

```

```

/plasmid="plasmid pGEM"
1
/misc_feature
/note="7 transcription initiation site"
11..44
/misc_feature
/note="multiple cloning sites"
47..69
/promoter
/note="SP6 promoter"
52
/misc_feature
/note="SP6 transcription initiation site"
76..323
/misc_feature
/note="lac operon sequence"
86..102
/misc_feature
/note="pUC/M13 reverse primer binding site"
complement(88..90)
/gene="lacZ"
complement(88..90)
/misc_feature
complement(88..90)
/feature="lac2"
/feature="lac2 start codon"
110..126
/misc_feature
/feature="lac operator"
complement(1247..1207)
/gene="bla"
complement(1247..1207)
/CDS
/feature="bla"
/codon_start=1
/transl_table=11
/product="Beta-lactamase"
/db_xref="PID:g58171"
/translation="MSIQHFRVALIPFAECLPFAHPETLVKVKDAEDQLGARVY
IELDNSKKILESFPERFPMSTFKVLGHTLLSPIDAGAEQGLPFAFYSNDLVE
YSPVTEKHLTDGMTVRELCSAITMSDNTAANLLITIGPKELTALHNKGDRVPL
DRWPELNEATPNDPRDTPVAMPPTTLRKLTLGELLTLASRQOLIDWMEADKVGPL
LPSALPAGWFLADKSGAERSGKIIIAALGPGRKPSKIVLIYITSSVAIMERNKQA
ETGASLIRKM"
2546..3001
/misc_feature
/feature="phage fl region"
3002..3162
/misc_feature
/feature="lac operon sequence"
3122..3138
/misc_feature
/feature="pUC/M13 forward primer binding site"
3165..3181
/promoter
/feature="T7 promoter"

BASE COUNT 816 a 783 c 814 g 768 t
ORIGIN

Query Match 43.0%; Score 2200; DB 32; Length 3181,
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2200; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

Db 39 AGCTTGAGTATTCTATAGTGTACCTAAATAGCTTGGCGTAATCATGTGTCATAGCTGTT 98
Cp 5102 AGCTTGAGTATTCTATAGTGTACCTAAATAGCTTGGCGTAATCATGTGTCATAGCTGTT 5043
Db 99 CCGTGTAATTTGTTATCGCTCAACATTCACACAAATACAGACACACAGCAGCAGCAGCAG 158
Cp 5042 CCGTGTAATTTGTTATCGCTCAACATTCACACAAATACAGACACACAGCAGCAGCAGCAG 4983
Db 159 TGTAAAGCCCTGGGTCCTAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 218
Cp 4982 TGTAAAGCCCTGGGTCCTAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 4923
Db 219 CCGCTTTCAGTCGGAACCTGTGTGCGAGCTGATTAATGAATGAGTGAAGTGAAGTGAAGT 278
Cp 4922 CCGCTTTCAGTCGGAACCTGTGTGCGAGCTGATTAATGAATGAGTGAAGTGAAGTGAAGT 4863
Db 279 GCGAGAGCGGTTTGGGTATTGGGCGCTCTTCGCTTCTCTGCTCACTCACTCACTCACT 338
Cp 4862 GCGAGAGCGGTTTGGGTATTGGGCGCTCTTCGCTTCTCTGCTCACTCACTCACTCACT 4803
Db 339 TCGGTCTGCTGCGTGGGCGAGGCTATGAGTCACTCAAGCGGTAATCGGTATCC 398
Cp 4802 TCGGTCTGCTGCGTGGGCGAGGCTATGAGTCACTCAAGCGGTAATCGGTATCC 4743

```



```

REFERENCE 3 (bases 1 to 3199)
AUTHORS Klock C.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-1993) Technical Services, Promega Corporation,
2800 Woods Hollow Road, Madison, WI 53711-5399, USA
COMMENT This vector can be obtained from Promega Corporation, Madison, WI.
Call one of the following numbers for order or technical
information:
Order or Technical 800-356-9526
In Wisconsin 800-356-9526
Outside U.S. 608-274-4330
See <X65300-X65335> for related vector sequences.
FEATURES
source
Location/Qualifiers
1..3199
/organism="Cloning vector"
/plasmid="Plasmid pGEM"
misc_feature
1
/Note="T7 transcription initiation site"
misc_feature
5..61
/Note="multiple cloning sites"
promoter
64..86
/Note="SP6 promoter"
misc_feature
69
/Note="SP6 transcription initiation site"
misc_feature
94..323
complement(106..108)
/Note="lac operon sequence"
misc_feature
/Note="lacZ"
/Note="lacZ start codon"
gene
complement(106..108)
/Note="lacZ"
gene
complement(1265..2125)
/Note="bla"
CDS
complement(1265..2125)
/Note="bla"
/codon_start=1
/transl_table=11
/product="Beta-lactamase"
/db_xref="PID:g58173"
/translation="MSIQHFRVALIPFFAAFLPVFAHPETLVKVKDAEDQLGARVGY
IELDLNGKILLESFRPEERPMWSTFKVLLCHTLISRIDAGDQELGRARYSONDLVE
YSPVTEKHLLDTGKTVBELCSAALTNSDNTAANLLITITGGKKEULTAFHNGDVRTEL
DRPELNEAIPNDERDITMPVAMPITLLRKLITGLLELLASRQQLIDWMEADKRVAGPL
LRSLALPAGWFIADKSGAGERGSPGIIAAIGPDGKPSRIIVYITGSGQATMDERNPQIA
EIGASLIKHW"
misc_feature
2564..3019
/Note="phage f1 region"
misc_feature
3020..3190
/Note="lac operon sequence"
promoter
3183..3199
/Note="T7 promoter"
BASE COUNT 779 a 803 c 801 g 816 t
ORIGIN
Query Match 43.0% Score 2200; DB J2; Length 3199.
Seq from Similarity 100.0% Pred. No. 0.00e+00;
Matches 2200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 57 AGCTTGAGTATCTATAGTGCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTT 116
Cp 5102 AGCTTGAGTATCTATAGTGCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTT 5043
Db 117 CTGTGTGAATTTATTCGCTCACAATTCACACACATACAGCGCGAAGCATAAAG 176
Cp 5042 CCGTGTGAATTTATTCGCTCACAATTCACACACATACAGCGCGAAGCATAAAG 4983
Db 177 TGTAAAGCTGGGTGCTAATAGTGAAGTACATCATATATTCGGTTCGCTCACTG 236
Cp 4982 TGTAAAGCTGGGTGCTAATAGTGAAGTACATCATATATTCGGTTCGCTCACTG 4923
Db 237 CCGGTTCCTCAATGCGGAAACCTGTGTGCGACGTGCAATTAAATGAATCGGCCAACCGCG 296
Cp 4922 CCGGTTCCTCAATGCGGAAACCTGTGTGCGACGTGCAATTAAATGAATCGGCCAACCGCG 4863

```

```

Db 297 GGAGAGGGGGTTTGGGTATTTGGGGCTCTCTCCGCTTCTCTGCTCAGTCTCGGTGGC 356
Cp 4862 GGAGAGGGGGTTTGGGTATTTGGGGCTCTCTCCGCTTCTCTGCTCAGTCTCGGTGGC 4903
Db 357 TCGTTCGTTTCGGCTGCGGGGAGAGGATATAGTCTCAATCAAAAGGAGGATATATTC 416
Cp 4802 TCGTTCGTTTCGGCTGCGGGGAGAGGATATAGTCTCAATCAAAAGGAGGATATATTC 4743
Db 417 ACAGATCAGGGGATATACGACGAGAAAGACATGTGAGCAAAAGGAGGAGGAGGAGG 476
Cp 4742 ACAGATCAGGGGATATACGACGAGAAAGACATGTGAGCAAAAGGAGGAGGAGGAGG 4663
Db 477 AACCGTAAAAAGGCGCGTTTGGGTATTTGGGTATTTGGGTATTTGGGTATTTGGGTAT 536
Cp 4682 AACCGTAAAAAGGCGCGTTTGGGTATTTGGGTATTTGGGTATTTGGGTATTTGGGTAT 4623
Db 537 CACAAAATCGAGCTCAAGTCAAGTCAAGTGGCGAAACCGGACAGGACTATAAGATACAG 596
Cp 4622 CACAAAATCGAGCTCAAGTCAAGTCAAGTGGCGAAACCGGACAGGACTATAAGATACAG 4563
Db 597 GCGTTTCCCGCTGGAAGCTCCCTGGTGGGCTCTCTCTGCTCAGTCTCGGTGGC 656
Cp 4562 GCGTTTCCCGCTGGAAGCTCCCTGGTGGGCTCTCTCTGCTCAGTCTCGGTGGC 4503
Db 657 TACGTCTCCCGCTTCTCTGCTGGGAAAGCGTGGGCTCTCTCTGCTCAGTCTCGGTGGC 716
Cp 4502 TACGTCTCCCGCTTCTCTGCTGGGAAAGCGTGGGCTCTCTCTGCTCAGTCTCGGTGGC 4443
Db 717 TATCTCAGTTCGGGTGAGTTCGCTCCAGCTGGGCTGTGTGTGAGGAAACCGGCTGT 776
Cp 4442 TATCTCAGTTCGGGTGAGTTCGCTCCAGCTGGGCTGTGTGTGAGGAAACCGGCTGT 4383
Db 777 CAGCGGAGCGCTCGGCTTATCCGCTTACCTGCTCTCTGCTGCTGCTGCTGCTGCTGCT 836
Cp 4382 CAGCGGAGCGCTCGGCTTATCCGCTTACCTGCTCTCTGCTGCTGCTGCTGCTGCTGCT 4323
Db 837 GACTTATCGCACTGGCAGCAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 896
Cp 4322 GACTTATCGCACTGGCAGCAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4263
Db 897 GGTGCTACAGAGTCTTGAAGTGGTGGCTTAACTAGGCTACACTAGAGGAGGACATATT 956
Cp 4262 GGTGCTACAGAGTCTTGAAGTGGTGGCTTAACTAGGCTACACTAGAGGAGGACATATT 4203
Db 957 GGTATCTGCGCTCTGCTGAGGCGAGTACCTGCGGAAAGAGATGGTAGCTCTGATGCT 1016
Cp 4202 GGTATCTGCGCTCTGCTGAGGCGAGTACCTGCGGAAAGAGATGGTAGCTCTGATGCT 4143
Db 1017 GGCAACAAACACCGCTGGTACGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1076
Cp 4142 GGCAACAAACACCGCTGGTACGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4083
Db 1077 AGAAAAAAGGATCTCAAGAGATCCTTTGATCTTTTCTAGGCGGTCTTACGCTCAGTGG 1136
Cp 4082 AGAAAAAAGGATCTCAAGAGATCCTTTGATCTTTTCTAGGCGGTCTTACGCTCAGTGG 4023
Db 1137 AACGAAACTCAGTTAAGGATTTTGGTTCATGAGATTAACAAAGAGATCTTACGCTCAG 1196
Cp 4022 AACGAAACTCAGTTAAGGATTTTGGTTCATGAGATTAACAAAGAGATCTTACGCTCAG 3963
Db 1197 ATCCTTTTAAATTAAGATTTTAAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1256
Cp 3962 ATCCTTTTAAATTAAGATTTTAAATCAATCAATCAATCAATCAATCAATCAATCAATCA 3903
Db 1257 TCTCAGAGTACCAATGCTTTAATCAGTGGGAGCTTATCTCAGGAGATCTTCTATTTCT 1316
Cp 3902 TCTCAGAGTACCAATGCTTTAATCAGTGGGAGCTTATCTCAGGAGATCTTCTATTTCT 3843
Db 1317 TCATCCATAGTTCGCTGACTCCCGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1376
Cp 3842 TCATCCATAGTTCGCTGACTCCCGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3783

```











Ds 1221 ATCTTTTAAATTAATAAGTTTAAATCAATCTAAAGTATATATAGTAAACTTGG 1280  
Cp 3962 ATCTTTTAAATTAATAAGTTTAAATCAATCTAAAGTATATATAGTAAACTTGG 3903  
Ds 1281 TCTGACAGTTACCAATGCTTAATCAGTGGGACCTATCTCAGCGATCTGTCTATTGGT 1340  
Cp 3962 TCTGACAGTTACCAATGCTTAATCAGTGGGACCTATCTCAGCGATCTGTCTATTGGT 3843  
Ds 1341 TCATCCATAGTTGCTGACTCCCGCTGCTGATAGTAACCTACGATACGGGAGGCTTACCA 1400  
Cp 3842 TCATCCATAGTTGCTGACTCCCGCTGCTGATAGTAACCTACGATACGGGAGGCTTACCA 3783  
Ds 1401 TCTGACAGTTACCAATGCTTAATCAGTGGGACCTATCTCAGCGATCTGTCTATTGGT 1460  
Cp 3782 TCTGACAGTTACCAATGCTTAATCAGTGGGACCTATCTCAGCGATCTGTCTATTGGT 3723  
Ds 1461 GCAATAAACAGCAGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 1520  
Cp 3722 GCAATAAACAGCAGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 3663  
Ds 1521 TCCATCCAGTCTATTAAATTTGTTCCGCGGAGCTAGAGTAGTTCGCGAGTTAATAGT 1580  
Cp 3662 TCCATCCAGTCTATTAAATTTGTTCCGCGGAGCTAGAGTAGTTCGCGAGTTAATAGT 3603  
Ds 1581 TTGGCAACGTTGTTGCGATTTGCTACAGGATCTGTTGCTGACGCTGCTGCTGCTGATG 1640  
Cp 3602 TTGGCAACGTTGTTGCGATTTGCTACAGGATCTGTTGCTGACGCTGCTGCTGCTGATG 3543  
Ds 1641 GCCTTCATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1700  
Cp 3542 GCCTTCATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3483  
Ds 1701 AAAAAAGCGTTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1760  
Cp 3482 AAAAAAGCGTTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3423  
Ds 1761 TTATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1820  
Cp 3422 TTATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3363  
Ds 1821 TGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1880  
Cp 3362 TGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3303  
Ds 1881 CCGAGTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1940  
Cp 3302 CCGAGTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3243  
Ds 1941 AAGTGTCTCATTTGGAACGCTTCTTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2000  
Cp 3242 AAGTGTCTCATTTGGAACGCTTCTTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3183  
Ds 2001 TTGAGTCCAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2060  
Cp 3182 TTGAGTCCAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3123  
Ds 2061 TTCACGAGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2120  
Cp 3122 TTCACGAGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3063  
Ds 2121 AGCGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2180  
Cp 3062 AGCGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3003  
Ds 2181 TATCAGGTTTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2240  
Cp 3002 TATCAGGTTTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2943  
Ds 2241 ATAGGAGTTTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3280  
Cp 2942 ATAGGAGTTTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3203

## RESULT 13

LOCUS CVGEM112P 3223 bp DNA SYN 12-FEB-1996  
DEFINITION Cloning vector pGEM-112f(+).  
ACCESSION X65313  
NID 958168  
KEYWORDS beta-lactamase; bla gene; cloning vector; lacZ gene; multiple cloning site; phage f1 region; promoter; pUC/M13 primer binding site.  
SOURCE unidentified cloning vector.  
ORGANISM unidentified cloning vector  
DEFINITION artificial sequence; cloning vectors.  
REFERENCE 1 (bases 1 to 3223)  
AUTHORS Technical Services.  
TITLE Direct Submission  
JOURNAL Submitted (23-MAY-1992) Technical Services, Promega Corporation.  
REMARK 2800 Woods Hollow Road, Madison, WI 53711-5399, USA  
REFERENCE 2 (bases 1 to 3223)  
AUTHORS Klock, C.  
TITLE Direct Submission  
JOURNAL Submitted (28-MAY-1993) Technical Services, Promega Corporation.  
COMMENT 2800 Woods Hollow Road, Madison, WI 53711-5399, USA  
This vector can be obtained from Promega Corporation, Madison, WI.  
Call one of the following numbers for order or technical information:  
Order or Technical 800-356-9526  
In Wisconsin 800-356-9526  
Outside U.S. 608-274-4330  
See <X65300-X65335> for related vector sequences.  
FEATURES  
source  
1..3223  
/organism="Cloning vector"  
/plasmid="plasmid pGEM"  
misc\_feature  
1  
/note="T7 transcription initiation site"  
misc\_feature  
19..85  
/note="multiple cloning sites"  
promoter  
88..110  
/note="SP6 promoter"  
misc\_feature  
93  
/note="SP6 transcription initiation site"  
misc\_feature  
118..347  
/note="lac operon sequence"  
misc\_feature  
128..144  
/note="pUC/M13 reverse primer binding site"  
gene  
complement(130..132)  
/gene="lacZ"  
misc\_feature  
complement(130..132)  
/gene="lacZ"  
misc\_feature  
152..168  
/note="lac start codon"  
gene  
complement(1289..2149)  
/gene="bla"  
CDS  
complement(1289..2149)  
/gene="bla"  
/codon\_start=1  
/transl\_table=11  
/product="beta-lactamase"  
/db\_xref="PID:958169"  
/translation="MSIQHFRVALIPFFAFCPLVFAHPETLVKVKDAEDQLARVGY  
IELDNSKILESFPPEPFPMSTFKVLLCHLLSIDAGDQGLPFPYSDNLI  
VSPVTEKHLTDGMTVRELSAAITMSDNTAANLLITIGGPKELTFLHMKGVHVR  
DRWPELNEATPNDERTDTPVAMPTTLIRKLTGELTLASROOLIDMKVAKVAPL  
LPSALPASKEFALAKSSAEPASPSIIAIAPGCGFSPVIVITIGSSATMEERKQTA  
EIGALPKLHW"  
misc\_feature  
2588..3043  
/note="phage f1 region"  
misc\_feature  
3044..3204  
/note="lac operon sequence"  
misc\_feature  
3164..3180  
/note="pUC/M13 forward primer binding site"  
promoter  
3207..3223

BASE COUNT	827 a	793 c	927 g	776 t	/note="T7 promoter"
ORIGIN					
Query Match	43.0%	Score 3209:	DB 32:	Length 3233:	
Best Local Similarity for 0k	Pr=4	N=0	0k=0.00		
Matches 2200:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:	
Db	81	AGCTTCAGTATTCTATATCTGCTACCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	140		
Cp	5102	AGCTTCAGTATTCTATAGTCTGCTACCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	5043		
Db	141	CGCTGTGAAATGCTATCTGCTACCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	200		
Cp	3642	CGCTGTGAAATGCTATCTGCTACCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	4983		
Db	261	TTTAAATGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	260		
Cp	4982	TTTAAATGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	4923		
Db	261	CGCTGTGAAATGCTATCTGCTACCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	320		
Cp	4922	CGCTGTGAAATGCTATCTGCTACCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	4863		
Db	321	GGCAGCTGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	380		
Cp	4862	GGCAGCTGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	4803		
Db	381	TGCTGTGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	440		
Cp	4802	TGCTGTGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	4743		
Db	441	ACAGAAATGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	500		
Cp	4742	ACAGAAATGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	4683		
Db	501	AGCTTTAAATGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	560		
Cp	4582	AGCTTTAAATGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	4623		
Db	561	CACAAATGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	620		
Cp	4522	CACAAATGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	4563		
Db	621	GGCTTCCGCTGGAAATGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	680		
Cp	4562	GGCTTCCGCTGGAAATGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	4503		
Db	681	TAGCTTGGCTTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	740		
Cp	4502	TAGCTTGGCTTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	4443		
Db	741	TATCTAGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	800		
Cp	4442	TATCTAGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	4383		
Db	801	CAGCTGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	860		
Cp	4382	CAGCTGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	4323		
Db	861	GACTTATGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	920		
Cp	4322	GACTTATGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	4263		
Db	921	GCTGTGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	980		
Cp	4262	GCTGTGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	4203		
Db	981	GGTATGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	1040		
Cp	4202	GGTATGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	4143		
Db	1041	GGTAAATGCTTGGGTGCTAAATAGCTTTGGGTAAATATATGTCATATAGCTGTTT	1100		

[illegible]

Db 2271 CACAAAATCGACGCTCAAGTCACAGTGGGCGAAAGCGGACAGGACATATAAAGATACAG 2330  
 Cp 4522 CACAAAATCGACGCTCAAGTCACAGTGGGCGAAAGCGGACAGGACATATAAAGATACAG 4563  
 Db 2331 CATTTCGCGCTGGAAGTCTCCCTGAGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2390  
 Cp 4592 CATTTCGCGCTGGAAGTCTCCCTGAGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4593  
 Db 2491 TACTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2450  
 Cp 4592 TACTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4443  
 Db 2451 TATCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2510  
 Cp 4442 TATCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4383  
 Db 2511 CAGCGGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2570  
 Cp 4382 CAGCGGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4323  
 Db 2571 GACTTATCGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2630  
 Cp 4322 GACTTATCGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4263  
 Db 2631 GGTGTACAGCTTCTGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2690  
 Cp 4262 GGTGTACAGCTTCTGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4203  
 Db 2691 GGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2750  
 Cp 4202 GGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4143  
 Db 2751 GCGAAGCAACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2810  
 Cp 4142 GCGAAGCAACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4083  
 Db 2811 AAAAAAAGATCTCAACAGAGATCTTGTGATCTTGTGATCTTGTGATCTTGTGATCTTGTGATCT 2870  
 Cp 4082 AAAAAAAGATCTCAACAGAGATCTTGTGATCTTGTGATCTTGTGATCTTGTGATCTTGTGAT 4023  
 Db 2871 AAGGAAGCTCAGCTTAAAGGATCTTGTGATCTTGTGATCTTGTGATCTTGTGATCTTGTGAT 2930  
 Cp 4022 AAGGAAGCTCAGCTTAAAGGATCTTGTGATCTTGTGATCTTGTGATCTTGTGATCTTGTGAT 3963  
 Db 2931 ATGCTTTAAATATAAATGAACTTTAAATCAATTAATGATATATGATGATGATGATGATGATGAT 2990  
 Cp 3962 ATGCTTTAAATATAAATGAACTTTAAATCAATTAATGATATATGATGATGATGATGATGATGAT 3903  
 Db 2991 TGTGACATTAATCAATGTTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3050  
 Cp 3902 TGTGACATTAATCAATGTTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3843  
 Db 3051 TATATCATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3110  
 Cp 3842 TATATCATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3783  
 Db 3111 TGTGCT 3170  
 Cp 3782 TGTGCT 3723  
 Db 3171 GCAATAAATCAGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3230  
 Cp 3722 GCAATAAATCAGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3663  
 Db 3231 TGTATCATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3290  
 Cp 3662 TGTATCATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3603  
 Db 3291 TGTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3350  
 Cp 3602 TGTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3543  
 Db 3351 GCTTATCATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3410

Cp 3542 GCTTATCATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4494  
 Db 3411 AAAAAAGCGGTAACTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4470  
 Cp 3482 AAAAAAGCGGTAACTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4424  
 Db 3471 TTATCAGCTCATGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4506  
 Cp 3422 TTATCAGCTCATGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4564  
 Db 3531 TGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4590  
 Cp 3362 TGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4564  
 Db 3591 CCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4606  
 Cp 3302 CCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4244  
 Db 3651 AAGTGTCTCATGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4710  
 Cp 3242 AAGTGTCTCATGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4194  
 Db 3711 TTGATCTCAGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4776  
 Cp 3182 TTGATCTCAGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4124  
 Db 3771 TGCACAGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4800  
 Cp 3122 TGCACAGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4764  
 Db 3831 AGGCGCAGCGGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4900  
 Cp 3062 AGGCGCAGCGGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4904  
 Db 3891 TATCAGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4900  
 Cp 3062 TATCAGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4904  
 Db 3951 ATAGCGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4990  
 Cp 2942 ATAGCGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2964

RESULT 15 CYP6M3Z 2743 bp DNA SYN 12 FEB 1996  
 LOCUS Cloning vector pGEW-42.  
 DEFINITION X65304  
 ACCESSION g58201  
 NID  
 KEYWORDS beta-lactamase; bla gene; cloning vector; lacZ gene; multiple cloning site; promoter.  
 SOURCE unidentified cloning vector.  
 ORGANISM artificial cloning vector.  
 REFERENCE 1 (bases 1 to 2743)  
 AUTHORS Technical Services  
 TITLE Direct Submission  
 JOURNAL Submitted (23-MAY-1992) Technical Services, Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711-5499, USA  
 REMARK revised by [2]  
 REFERENCE 2 (bases 1 to 2743)  
 AUTHORS Klock, C.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-MAY-1994) Technical Services, Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711-5499, USA  
 COMMENT This vector can be obtained from Promega Corporation, Madison, WI. Call one of the following numbers for order or technical information:  
 Order or Technical 800-459-9626  
 In Wisconsin 800-456-9526  
 Outside U.S. 608-274-4499  
 See X65300 for related sequences.  
 FEATURES Location/Qualifiers







\*\*\*\*\*  
 W P R E H  
 \*\*\*\*\*

(TW)

Release 3.1A John F. Collins, BioComputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPsearch\_n n.a. n.a. database search, using Smith-Waterman algorithm  
 Run on: Mon Mar 15 17:10:57 1999; MasPar time 725.31 Seconds  
 Tabular output not generated 959.461 Million cell updates/sec  
 Title: >US-09-020-716-5  
 Description: (1-5115) from US09020716.seq  
 Perfect Score: 5115  
 N.A. Sequence: 1 GTTGGGAGCTCTCCCATATG TCAAGCTATGCACTCAAGC 5115  
 Comp. CAACCTCGAGAGGATATAC AGTTCGATACGTAGGTGG

Scoring table: TABLE default  
 Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 58026449 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

n-gene32  
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
 14:part14 15:part15 16:part16 17:part17 18:part18  
 19:part19 20:part20 21:part21 22:part22 23:part23  
 24:part24 25:part25 26:part26 27:part27 28:part28  
 29:part29 30:part30 31:part31 32:part32 33:part33  
 34:part34 35:part35 36:part36 37:part37 38:part38  
 39:part39 40:part40

Statistics: Mean 10.354; Variance 5.797; scale 1.786

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2195	42.9	10306	12	Retro virus vector p5	0.00e+00
2	2195	42.9	10470	12	Retro virus vector p5	0.00e+00
3	2172	42.5	4118	35	Construct pGEM-HTR(KA)	0.00e+00
4	2172	42.5	4118	35	Construct pGEM-HTR(KA)	0.00e+00
5	2176	42.5	4283	35	DNA encoding hGH, EGF	0.00e+00
6	2176	42.5	4283	35	DNA encoding hGH, EGF	0.00e+00
7	2142	41.9	3690	39	Plasmid pBSGFP expres	0.00e+00
8	2142	41.9	4534	14	Plasmid pBSGFP expres	0.00e+00
9	2142	41.9	6295	39	Plasmid pWRG3196 enco	0.00e+00
10	2141	41.9	6824	6	K lactis/S. cerevisiae	0.00e+00
11	2141	41.9	7287	39	Plasmid pWRG3169 enco	0.00e+00
12	2142	41.9	10930	20	Plasmid pM16.1	0.00e+00
13	2142	41.9	10950	20	Plasmid pM16.1	0.00e+00

ID	Q73734	standard: DNA; 10306 BP.	Q43379	Sequence of clone ES2	0.00e+00
AC	Q73734		4145	Human cyclin D1-human	0.00e+00
DE	Q73734		4540	Shortened C1 gene in	0.00e+00
DE	Q73734		4824	pDEL10.	0.00e+00
KW	Q73734		4883	PrEt-Splice.	0.00e+00
KW	Q73734		5178	Plasmid pRIPPHAT (rat	0.00e+00
KS	Q73734		5356	PUMIT sequence incl	0.00e+00
PD	Q73734		5534	Plasmid pMP6-112.	0.00e+00
PD	Q73734		5585	Plasmid pRC/CMV-HBS c	0.00e+00
PD	Q73734		5618	PVF108	0.00e+00
PD	Q73734		5620	RSV tar Rev M10 expre	0.00e+00
PD	Q73734		5653	pRL2034 construct.	0.00e+00
PD	Q73734		6207	Plasmid pRC/ASK8-SmGS	0.00e+00
PD	Q73734		6216	Beta-lactamase vector	0.00e+00
PD	Q73734		6230	PCMV-BL for extracell	0.00e+00
PD	Q73734		6231	Plasmid pFRED7 expres	0.00e+00
PD	Q73734		6238	pRLD203 Apal constr	0.00e+00
PD	Q73734		6274	Vector pCDNA3-IgG1.	0.00e+00
PD	Q73734		6338	Vector COSpLink	0.00e+00
PD	Q73734		6367	Vector COSpLink cont	0.00e+00
PD	Q73734		6367	Plasmid pCMV-BL10M en	0.00e+00
PD	Q73734		6555	IL-4.Y124D/IgM1 fusio	0.00e+00
PD	Q73734		6926	Sequence of pIE4.	0.00e+00
PD	Q73734		7050	PROV167 sequence	0.00e+00
PD	Q73734		7106	Plasmid pRCV-167 codi	0.00e+00
PD	Q73734		7106	Plasmid pRCV-162 codi	0.00e+00
PD	Q73734		7298	Plasmid pR73 encodin	0.00e+00
PD	Q73734		7474	PSC11 FIPV EL.	0.00e+00
PD	Q73734		8710	PSC11 FIPV N.	0.00e+00
PD	Q73734		9020	Plasmid pCB51 encodin	0.00e+00
PD	Q73734		10288	Plasmid pR72 encodin	0.00e+00
PD	Q73734		10443	Plasmid pR72 encodin	0.00e+00
PD	Q73734		10443	Eukaryotic layered ve	0.00e+00
PD	Q73734		16656		0.00e+00

## ALIGNMENTS

RESULT 1  
 ID Q73734 standard: DNA; 10306 BP.  
 AC Q73734;  
 DE Retro virus vector p521.  
 KW GalV, gibbon ape-leukemia virus, retro virus, vector, plasmid p558;  
 KS plasmid p521; plasmid p537; transgenic animal; gene therapy; ss.  
 SS Synthetic.  
 PD W09423048-A.  
 PD 13-OCT-1994.  
 PD 06-APR-1994; 003784.  
 PR 06-APR-1993; US-043311.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PI Deacon NJ, Eiden MV, Hooker DJ, Wilson CA;  
 DR WPI: 94-333206/41.  
 PT Recombinant DNA contg. defective gibbon ape leukaemia virus  
 PT genome - and specific gene, also derived mammalian cells, viruses  
 PT etc., useful in gene therapy of e.g. cancer, viral infections or  
 PT inherited diseases  
 PS Disclosure: Page 41-46; 68pp; English.  
 CC Novel replication-deficient retro virus vectors p558 (Q73737), p521  
 CC (Q73734) and p537 (Q73735) include a GalV-derived packaging site,  
 CC esp. the 200-910 base region of the GalV SEA0 genome (Q73731), and  
 CC regulatory sequences, partic. the GalV 3' long terminal repeat  
 CC (Q73732). The vectors allow transfer of genes for gene therapy and  
 CC transgenic animal breeding.  
 SQ Sequence 10306 BP; 2400 A; 2758 C; 2733 G; 2415 T;  
 Query Match 42.9%; Score 2195; ES 12; Length 10306;  
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
 Matches 2195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 8112 cagtgagcacttttcagggaatgtatcgacgaacccatttttttttttataaac 8171  
 QY 2903 CAGTGGACATTTTCGGGAAATGTCGCGGACCCGATTTTCTTAAATAC 2962











ID	T86449 standard: DNA: 4283 BP.		
AC	185449:		
DI	10-FEB-1998	(first entry)	
DE	DNA encoding hGH, EGF fusion protein from pWRG1630.		
KW	Mature epidermal growth factor; EGF; wound healing; skin, gold bead;		
KW	epidermal cell; hair follicle; polyA addition sequence;		
KW	human growth hormone secretory signal peptide; hGH; ds.		
OS	Chimeric - homo sapiens.		
OS	Chimeric - bos taurus.		
OS	Synthetic.		
FE	Key	Location/Qualifiers	
FE	CDS	713..1253	
FI	/note=	"contains intronic sequences"	
FI	/tag=	a	
FI	/note=	"contains intronic sequences"	
FI	/tag=	b	
FI	/standard_name=	secretion_signal	
FI	/note=	"human growth hormone secretory signal contains intronic sequences"	
FI	mat_peptide	1050..1250	
FI	/tag=	c	
FI	/product=	mature_human_EGF	
FI	/tag=	d	
FI	/number=	1	
FI	/tag=	e	
FI	/number=	1	
FI	/tag=	f	
FI	/number=	2	
FI	/note=	"end of exon is given as TAG stop codon"	
FI	US5661132-A.		
PD	26-AUG-1997.		
PF	22-NOV-1994:	343401.	
PR	14-DEC-1989:	US-451957.	
PP	22-MAY-1991:	US-7072348.	
PR	11-JUN-1992:	US-897357.	
PR	11-JUN-1993:	US-076550.	
PA	(AURA-) AURAGEN INC.		
PI	Andree C. Eriksson E. Macklin MD, Swain WF.		
PI	WPI: 97-434422/40.		
DR	P-PSDB: W32221.		
PT	Genetic construct encoding mature epidermal growth factor - when delivered to cells on gold beads is useful for treatment of skin wounds		
PS	Claim 3: Columns 11-18, 15pp. English		
CC	This sequence represents a new genetic construct consisting of, in 5'-to-3' order, a promoter effective in skin cells; a signal sequence that directs secretion of a protein from a skin cell; a non-naturally occurring DNA fragment which encodes an epidermal growth factor (EGF) and lacks all coding regions of a natural EGF precursor gene except the coding region that encodes the mature EGF; and a polyA addition sequence. The genetic construct is used in a method for treating a skin wound site on an animal. The presence of the EGF accelerates, by a statistically significant amount, the healing time of a wound treated in this manner. Also the need to use expensive purified mature EGF protein is avoided.		
CC	Sequence	4283 BP: 1023 A; 1095 C; 1114 G; 1051 T;	
SQ			
Query Match 42.5%; Score 2176. DB 35. Length 4283;			
Best Local Similarity 99.5%; Pred No 0.00e+00;			
Matches 2198; Conservative 0; Mismatches 12; Indels 0; Gaps 0;			
Db	1597	agnttaagattctatagtgccactaaatagctggcgataatcgtatagctgttt	1656
Cp	5102	AGGTTCAGTATTCATAGTGTACCTAAATAGCTGGGCGTAATCATAGTCIGTIT	5043
Db	1657	ccgtgtgaaatggtatccgctccaaattccacacacatacagccggaagcataaag	1716
Cp	5042	CTCTGTGCAATTGTATTGCGGTCAATTCACACACATACAGACCGGAAGCATAAAG	4993
Db	1717	tataaagccctggggtgcctaataatgagtgagctcaactcacattaattgcttgcgtccactg	1776







Dh	2137	gactttcccccggaaagctccctctatggctctctctgttccgaacctgcgcgttacccga	2196
Cp	4562	ggctttcccccctgaaagctccctctgtgctgctctgttccgaacctgcgcgttacccga	4503
Dh	2197	tacctatcgcccttccctctcggaaagcctgagcgctttctcaagctcaacctgtagg	2256
Cp	4502	tacctatcgcccttccctctcggaaagcctgagcgctttctcaagctcaacctgtagg	4443
Dh	2257	tactcaagttccggtgtagctgttcgcctccaagctggcgctgtgacagcaaccccccggt	2316
Cp	4442	tatctcagttccggtgtagctgttcgcctccaagctggcgctgtgacagcaaccccccggt	4383
Dh	2317	caaccgacccgtgcgccttatccggttaactatcgtcttgagttcccaaccocggtaaacac	2376
Cp	4382	caaccgacccgtgcgccttatccggttaactatcgtcttgagttcccaaccocggtaaacac	4323
Dh	2377	gacttatgccactggcagcagccactggttaacaggattagcaagcagagttatgagc	2436
Cp	4322	gacttatgccactggcagcagccactggttaacaggattagcaagcagagttatgagc	4263
Dh	2437	gggtctacagagttcttgaagttggcctaactacgctacacataagaagcagattatt	2496
Cp	4262	gggtctacagagttcttgaagttggcctaactacgctacacataagaagcagattatt	4203
Dh	2497	ggatctgcgctctcgtcgaagccagttaccttcgaaagaaagtgtgtagctcttatcc	2556
Cp	4202	ggatctgcgctctcgtcgaagccagttaccttcgaaagaaagtgtgtagctcttatcc	4143
Dh	2557	ggcaacaacaaccacgctggtgagcgtagttttttgtttcaagcagcagattagcgc	2616
Cp	4142	ggcaacaacaaccacgctggtgagcgtagttttttgtttcaagcagcagattagcgc	4083
Dh	2617	agaaaaaagatctcaagaagatcctttgatctttctacggggtctgacgctcaatgg	2676
Cp	4082	agaaaaaagatctcaagaagatcctttgatctttctacggggtctgacgctcaatgg	4023
Dh	2677	aacgaaaactcacgttaagggaatttggctcatgaattatcaaaaagatcttcacctag	2736
Cp	4022	aacgaaaactcacgttaagggaatttggctcatgaattatcaaaaagatcttcacctag	3963
Dh	2737	atccttttaattaaaaatgaagttttaaatcaactcaaatatataatgagtaaaccttg	2796
Cp	3962	atccttttaattaaaaatgaagttttaaatcaactcaaatatataatgagtaaaccttg	3903
Dh	2797	tctgacagttaccaatgcttaatcagtaggacacattctcagcgatctgtctattcgt	2856
Cp	3902	tctgacagttaccaatgcttaatcagtaggacacattctcagcgatctgtctattcgt	3843
Dh	2857	tcatacagttaccctgactccccctcgtgtagataactacgatacgggagggcttacc	2916
Cp	3842	tcatacagttaccctgactccccctcgtgtagataactacgatacgggagggcttacc	3783
Dh	2917	tctgccccagtagtcaatgataccggcagacccacgctcacccggtccagattatca	2976
Cp	3782	tctgccccagtagtcaatgataccggcagacccacgctcacccggtccagattatca	3723
Dh	2977	gcaataaaccacgaccccggaagggccggagcacaagaatggtctgcacacttalcgc	3036
Cp	3722	gcaataaaccacgaccccggaagggccggagcacaagaatggtctgcacacttalcgc	3663
Dh	3037	tccatccagctcttaattggttcgcgggaagctagagtaagtgttcgcagattaagt	3096
Cp	3662	tccatccagctcttaattggttcgcgggaagctagagtaagtgttcgcagattaagt	3603
Dh	3097	ttgcgcacactgtttgcattgctacaggaactcgttgtgtcacgcgtcgcttgggtatg	3156
Cp	3602	ttgcgcacactgtttgcattgctacaggaactcgttgtgtcacgcgtcgcttgggtatg	3543
Dh	3157	gcttcattccagctcccggttcccacaacatcaaggcagattacatatcccccatgttgc	3216
Cp	3542	gcttcattccagctcccggttcccacaacatcaaggcagattacatatcccccatgttgc	3483

Db	3217	aaaaaacgggttagctccttcggctcccgatcgtttcacaagaatgagcgcgaatg	3276
Cp	3482	AAAAAAGCGGTTAGCTCCCTCGGTCCTCCGATCGTGTGTCAAGAAJAAJTTT	3423
Db	3277	ttatcactcatggttatgcacacactgcataattccttactctatgcacatccat	3336
Cp	3422	TTATCACTCATGGTTATGAGACCAATGATAATCTCTTAATGCTATGCTATG	3363
Db	3337	tgtttttctgtgactgtaagtactcaacccaagtcattctgaagataatgtatgcgcga	3396
Cp	3362	TGCTTTCTGTGACIGTGAGTAGTCAACCAAGTCAATCTGAGAAATACGCGCGCGG	3403
Db	3397	ccgagttgctcttgcgcgcgcgcataacacgggataataccggcgccacatagcagaacttta	3456
Cp	3302	CCGAGTTCTGCTTTGCGCGCGTCAATACCGGATATATAGTGTATGACATAGCAGAACTTAA	3243
Db	3457	aaagtcctcatcattgaaacagttctcttcggggcgaaacctcgaagatctttaccgctg	3516
Cp	3242	AAAGTGCTCATCTTGAAGAACGTTCTTTCGGGCGAAACATCTCAAGATCTTACCGCTG	3183
Db	3517	ttgaatccagttcctgagttaaccacactcgtgcaccccaactaatcttccagctcttttact	3576
Cp	3182	TTGAGATCCAGTTTCGATGTAAACCCACTCTGTGCACCAACTGATCTTCAGCATCTTTTACT	3123
Db	3577	ttaccacagcgtttctcgggtgagcaaaaacaggaagcaaaatgccgcaaaaagggaata	3636
Cp	3122	TTCAACAGCGTTTCTGGGTGACCAAAAACAGCAAGCGCAAAATGCGGCAAAAAGGGGATA	3083
Db	3637	aggcgacacgcgaatgttgataactcatcactctctcttcttccaatattattgaacatt	3696
Cp	3062	AGGTCGACAGCGAAATGTTGAATACTCATACTCTCTCTCTTTTTCATATATTGAAGCAT	3003
Db	3697	tatcaggttatgtctcatgacgagatcacattttgaatctatttcaaaaataaacaa	3756
Cp	3002	TATCAGGGTTATTGCTCATGACGGATACATATTGTAATGTATTAGAAAATAAACAA	2943
Db	3757	ataaggggttcgcgcacatttcccgaaaagtcgcacctg	3796
Cp	2942	ATAGGGGTTCGGCGCAATTTCGCGCAAAAGTGCACCTG	2903
RESULT 7			
ID	V14340 standard; DNA: 3699 BP.		
AC	V14340;		
DT	19-MAY-1998 (first entry)		
DE	Plasmid pBSGFP expressing green fluorescent protein.		
KW	Green fluorescent protein; GFP: blue fluorescent protein; BFP: jellyfish		
KW	genetic engineering marker; gene therapy; plasmid; ss.		
OS	Synthetic.		
PN	W059742320-A1.		
PD	13-NOV-1997.		
FF	07-MAY-1997; U07625.		
PR	08-MAY-1996; US-646538.		
PA	(USSH ) US SEC DEPT HEALTH.		
PA	(USSH ) US DEPT HEALTH & HUMAN SERVICES.		
F1	Gaitanaris GA, Pavlakis GN, Stauber RH, Vournakis JN;		
DR	WPI; 97-558982/51.		
CC	This sequence is a plasmid used to express the mutated green fluorescent		
CC	protein (GFP) of Aequorea victoria coding sequence of the invention. The		
CC	DNA of the invention that encodes a protein that is (a) a protein that has		
CC	has Leu at position 65 (and optionally Thr at 168 and optionally further		
CC	Cys at 66) and has cellular fluorescence at least 5 times that of		
CC	wild-type GFP or (b) is a blue fluorescent protein (BFP) with His at		
CC	position 67 and also at least one of Leu at 65 and Ala at 164, and has		
CC	cellular fluorescence at least 5 times that of BFP (Lyr67 to His). The		
CC	nucleic acids can be used as markers in genetic engineering and gene		
CC	therapy. They may also be used to detect and characterize regulatory		
CC	coding sequence elements that control subcellular expression and		

CC targeting of proteins. Typical applications are monitoring targeting and  
 CC transport of proteins in cells, assessment of gene therapy procedures, in  
 CC diagnosis (when expressed under control of a promoter induced by a  
 CC particular analyte), assessment of mutagenicity of compounds, and for  
 CC drug screening (where expression is controlled by the promoter of a  
 CC target gene), particularly for antiviral or antiparasitic agents  
 CC sequence 3694 bp; 942 A; 884 C; 898 G; 975 T;

Query Match 41.9%; Score 2142; DB 39; Length 3699;  
 Best local similarity 94.4%; Prod No 0 non-matching  
 Mismatches 2156; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

DB 1560 acattacataacacatcatatatttcttatataatattatccagctcacatt 1589  
 |||||  
 Cp 6372 AGTTGGGTAAATCAAGGTCAAGCTTCTCTCTGTAATTTTATCGCTCACATT 5013  
 |||||  
 Db 3640 caacacacataac 1649  
 |||||  
 Cp 6312 CCAACACATACAGATACATACATACATACATACATACATACATACATAC 4953  
 |||||  
 Db 1650 taactcaatataatataatataatataatataatataatataatataatata 1769  
 |||||  
 Cp 4952 TAAAT 4893  
 |||||  
 Db 1710 taactcaatataatataatataatataatataatataatataatataatata 1769  
 |||||  
 Cp 4842 CAGTGTATTAATCAATACATACATACATACATACATACATACATACAT 4833  
 |||||  
 Db 1770 taactcaatataatataatataatataatataatataatataatataatata 1829  
 |||||  
 Cp 4842 TGGGTCTTCT 4773  
 |||||  
 Db 1840 taactcaatataatataatataatataatataatataatataatataatata 1889  
 |||||  
 Cp 4772 TATCATCAAAAGCGGTAAATATATATATATATATATATATATATATAT 4713  
 |||||  
 Db 1960 atataac 1949  
 |||||  
 Cp 4712 AATGATCAAAAGCGGTAAATATATATATATATATATATATATATATAT 4653  
 |||||  
 Db 1960 taactcaatataatataatataatataatataatataatataatataatata 2009  
 |||||  
 Cp 4652 TGGGTCTTCT 4593  
 |||||  
 Db 2010 caac 2069  
 |||||  
 Cp 4592 CCAAAAGCGGTAAATATATATATATATATATATATATATATATATATAT 4533  
 |||||  
 Db 2070 taactcaatataatataatataatataatataatataatataatataatata 2129  
 |||||  
 Cp 4532 TGGGTCTTCT 4473  
 |||||  
 Db 2140 taactcaatataatataatataatataatataatataatataatataatata 2189  
 |||||  
 Cp 4472 CAGTGTATTAATCAATACATACATACATACATACATACATACATACAT 4413  
 |||||  
 Db 2190 taactcaatataatataatataatataatataatataatataatataatata 2249  
 |||||  
 Cp 4412 AAGCTGGGTCT 4353  
 |||||  
 Db 2250 taactcaatataatataatataatataatataatataatataatataatata 2309  
 |||||  
 Cp 4352 TATGTCTTCT 4293  
 |||||  
 Db 2410 taactcaatataatataatataatataatataatataatataatataatata 2369  
 |||||  
 Cp 4292 AATAGGATTAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG 4233  
 |||||  
 Db 2370 taactcaatataatataatataatataatataatataatataatataatata 2429  
 |||||  
 Cp 4232 AATATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG 4173  
 |||||  
 Db 2430 taactcaatataatataatataatataatataatataatataatataatata 2489

Cp 4172 TCGSAAAAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 4113  
 |||||  
 Db 2490 ttt 2449  
 |||||  
 Cp 4112 TTTTGTGTTTGAAAGCAGCAGATACATGATGATGATGATGATGATGATGAT 4053  
 |||||  
 Db 2550 atcttttttctacggggtctacggggtctacggggtctacggggtctacggggt 2499  
 |||||  
 Cp 4052 ATCTTTTCTAGGGGTTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 4003  
 |||||  
 Db 2610 atgagattatcaaaagagatcttccactccactccactccactccactccactcc 2569  
 |||||  
 Cp 3992 ATGAGATTATCAAAAGAGATCTTCAGCTAGATGCTTTTAATTTAAATTAAG 3933  
 |||||  
 Db 2670 taactcaatataatataatataatataatataatataatataatataatata 2629  
 |||||  
 Cp 3932 TCAATCTAAGAT 3873  
 |||||  
 Db 2730 gcacattctcagcgatctctctctctctctctctctctctctctctctctctct 2789  
 |||||  
 Cp 3872 GCACTATCTTCAAGGATATATATATATATATATATATATATATATATAT 3813  
 |||||  
 Db 2790 tagataactacgatacagggagggcttccactccactccactccactccactcc 2849  
 |||||  
 Cp 3812 TAGATAAATAGATATAGGAGATGAGGATGAGGATGAGGATGAGGATGAGG 3753  
 |||||  
 Db 2850 gacccacactcagcgatctctctctctctctctctctctctctctctctctctct 2909  
 |||||  
 Cp 3752 GACCCAGCTCTACGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3693  
 |||||  
 Db 2910 ccacagatggtctctcgaactttctccgctccactccactccactccactccact 2969  
 |||||  
 Cp 3692 CCACAAATGCT 3633  
 |||||  
 Db 2970 gctagagtaagtaattccgcaatttaatttttgcgaagcttatttgcgaagcttatt 3029  
 |||||  
 Cp 3632 GTAT 3573  
 |||||  
 Db 3030 atcttttttctacggggtctacggggtctacggggtctacggggtctacggggt 3089  
 |||||  
 Cp 3572 ATCTTTTCTAGGGGTTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 3513  
 |||||  
 Db 3090 agggaggttactatgacccctctctctcgaagagagagagagagagagagagagag 3149  
 |||||  
 Cp 3512 AGGCGAGTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3453  
 |||||  
 Db 3150 atcgttgtcagaagtaatttgcgcaattgttattccactccactccactccactcc 3209  
 |||||  
 Cp 3452 ATCTTTTCTAGGGGTTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 3393  
 |||||  
 Db 3210 aattcttctacgatacaggggtctacggggtctacggggtctacggggtctacggggt 3269  
 |||||  
 Cp 3392 AATCTCTTACTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 3333  
 |||||  
 Db 3270 aagtcattctcagagtaatttctctctctctctctctctctctctctctctctctct 3329  
 |||||  
 Cp 3332 AAGTCATTCTGAGAT 3273  
 |||||  
 Db 3330 gataac 3389  
 |||||  
 Cp 3272 GATAATAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3213  
 |||||  
 Db 3390 gggcgaacactctcagagtaatttctctctctctctctctctctctctctctctctctct 3449  
 |||||  
 Cp 3212 GGGGCAAAATCTCAAGATATATATATATATATATATATATATATATATATAT 3153  
 |||||  
 Db 3450 gcacccactaatctcagagtaatttctctctctctctctctctctctctctctctctctct 3509  
 |||||  
 Cp 3152 GCACCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3093  
 |||||  
 Db 3510 ggaagcgaacactcagagtaatttctctctctctctctctctctctctctctctctctct 3569  
 |||||





Cp 4712 ATGTGAGCAAAAGGCGAGCAAAAGGCGAGAAAGCGGAGTTGTTGGCGTTT 4653  
Db ttcacatagctccgccccctcaacagcatcacaaaaatcaacactcaagtcacagtg 4141  
Cp 4552 TTGATAGGCTGGCTCCCTTACGAGGATATCAAAAATGAGCGCTCAAGTCAGAGTG 4593  
Db cgaacccccacaggactataagataaccagggtttcccccctgaagctccctcgtgc 4201  
Cp 4592 CGAAACCCGACAGGACTATAAGATACAGAGGCTTTCCCCCTGGAAGCTCCCTGTCGC 4533  
Db tctcctgttccacccctgcgcttaccggatcacctgtccgcttctccttcgggaagc 4261  
Cp 4532 TCCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCTTCTCCCTTCGCGAAGC 4473  
Db gtggccttctcatagctccagctgttaggtatctcagttcgggttaggttcggtcc 4321  
Cp 4472 GTGGCGCTTTTCATAGCTTACGCTGTAGGTATCTCAGTTGGGTAGCTGTTGGTCC 4413  
Db aagctgggtgtgtgcacgaaccccccttccagcccgagcgtcgtcgttaccggtaac 4381  
Cp 4412 AAGCTGGGTGTGTGCACGAAGCCCGCTTCAGCCGAGCGCTGCGCTTATCCGGTAAAC 4353  
Db taccgtcttgagctccaaacccggtgaagacacgacttatcgccactggcagcagcactggt 4441  
Cp 4352 TATCGCTTTTASGTCTAACCCCGGTAAAGACACGACTTATGCGCACTGCGCAGCCACCTGGT 4293  
Db aacagattagcagagcaggtatgttaggggtgtgtacagatcttctgaagtggtgct 4501  
Cp 4292 AA-CAGGATTAGCAGAGCGAGGTATGAGCGGTGTCTACAGAGTCTTGAAGTGGTGGCT 4233  
Db aactacggttacactagaaggacagctatttggtatctcgtcgtcgtcgtgaagccagttacc 4561  
Cp 4232 AAGTAGGGTACACTASAGAGGACAGTATTGTGTATCTGCGCTCTGCTGAAGCCAGTTACC 4173  
Db ttcgaaaaagagttgtagctcttgatccgcgaacaaacacacgcgtgtagcgtggt 4621  
Cp 4172 TTGGGAAAAAGAGTTGTTAGTCTTGTATCCGGCAACCAACCCCGCTGGTAGCGGTGGT 4113  
Db tttttgtttgcaagcagcagatctacgcgcagaaaaaaaggatctcaagaagatcctttg 4681  
Cp 4112 TTTTITTTTAAAGTACAGATTATGCGCAGAAAAAAGGATCTCAAGAAGATCTTTTG 4053  
Db atcttttctacgggtctgaagctcagtgagaacgaacactcaagtcaggtatttggtc 4741  
Cp 4052 ATCTTTCTACGGGCTCTCAAGCTCAGTGTGAACGAAATCACTGTTAAGGATTTTGTC 3993  
Db atcagattatcaaaaagagcttccactagatccttttaatttaaaatgaagttttaa 4801  
Cp 3992 ATCAGATTATCAAAAAGGATCTTCACCTAGATCCCTTTTAAATTAATAATGAAGTTTAAA 3933  
Db tcaatctaaadtatatatagtaaaacttgctgacagttaccaatgcttaactcagtgag 4861  
Cp 3932 TCAATCTAAATATATATAGTAAGTAACTTGGTGTGACAGTTACCAATGCTTAAATCAGTGAG 3873  
Db gacactatctcagcagctctctattcttccatccatagttgctgactcccgctcgtg 4921  
Cp 3872 GCACCTATCTCAGCGATCTGTATTCTGTTATCTGTTATCTGTTGCTGACTCCCGTCTGT 3813  
Db tagtaactacgatacgggagggcttaccatctcgcccccauigtgcgaatgaataaccgga 4981  
Cp 3812 TAGATAACTATATAGGAGAGGCTTACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3753  
Db gacccagctccacgctccagatttatcagcaataaaacagccagccgagcggagccgag 5041  
Cp 3752 GACCCAGCTCAGCGCTCTCAGATTATCAGCAATAAAACAGCCAGCGGAAGGCGCGAG 3693  
Db cgcagaagtgctcctgcaactttatccgctccatccatccatctatttaattgttgcgggaa 5101  
Cp 3692 CGAGAGAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCTATTAAATGTTGGCGGAA 3633  
Db gctagaagtaagttgctgcccagttgaagtttgcgcaacggttggccactgctacagggc 5161  
Cp 3632 GCTAATA-TAAGTASGTAGTAAATATTTGGCGCAAGTGTGGGCACTGCTACAGGC 3573

## RESULT 10

ID Q39050 standard; DNA: 6824 BP.  
AC Q39050:  
DT 28-JUL-1993 (first entry)  
DE K.lactis/S. cerevisiae genetic vector.  
KW Genetic vector; integration; Kluyveromyces lactis; 25S ribosomal DNA;  
KW Saccharomyces cerevisiae; E. coli; domain; yeast; plasmid; promoter;  
KW expression cassette; HIS3; marker; transformant; human; lysozyme; HL2;  
KW GAL7; signal sequence; killer toxin; transcription termination signal;  
KW FLP; 2 micron plasmid; ss.  
OS Synthetic.  
PN EP-537456-A.  
PD 21-APR-1993.  
PF 31-AUG-1992: 114838.  
PR 04-SEP-1991: IT-M12349.  
PA (ISTS) SCLAVO SPA.  
PI Galicotti CU, Gallo E, Riccio ML, Rossolini GM, Thaller MC;  
DR WPI: 93-127394/16.  
PT Vector for Kluyveromyces lactis and Saccharomyces cerevisiae -  
PT heterologous proteins  
PS Claim 1: Fig 1: 26pp. English.  
CC This sequence represents a genetic vector which allows the stable



```

QY 4320 GTGCTGCTTACCGGCTTGGAGCTCAAGACGATAGTATTACGGGATAAGGCGCAGCGTCCGG 4379
Db 6026 CTGACGGGGGTTCTGACACAGCCAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6085
QY 4380 CTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4439
Db 6086 ATACCTACAGCTGAGCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6145
QY 4440 ATACCTACAGCTGAGCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4499
Db 6146 GTATCCGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6205
QY 4500 GTATCCGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4559
Db 6206 CAGCTGATCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6265
QY 4560 CAGCTGATCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4619
Db 6266 GTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6325
QY 4620 GTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4679
Db 6326 GTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6385
QY 4680 GTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4739
Db 6386 TATGATACCGTATTACCGCTTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6445
QY 4740 TATGATACCGTATTACCGCTTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4799
Db 6446 CAGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6505
QY 4800 CAGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4859
Db 6506 CAGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6565
QY 4860 CAGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4919
Db 6566 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6625
QY 4920 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4979
Db 6626 AGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6685
QY 4980 AGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5039
Db 6686 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6718
QY 5040 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5072

RESULT 11
ID V02042 standard. DNA. 7287 BP.
AC V02042;
DI 08-JUN-1998 (first entry)
DE Plasmid pWRG3169 encoding murine interleukin-12.
KW Interleukin-12; IL-12; cytokine; growth factor; mouse;
KW Plasmid pWRG3169; cancer; tumour; metastasis; gene therapy; ds;
KW Cyclical; Circular
KW Chimeric - Mus musculus.
OS Chimeric - Cytomegalovirus.
OS Chimeric - Bos taurus.
OS Chimeric - Rhesus macaque polyoma virus.
PH Key Location/Qualifiers
FT promoter 1..628
FT /tag= a
FT /note= "CMV promoter"
FT iDNA 629..810
FT /tag= b
FT CDS 953..1673
FT /tag= c
FT /product= p35 subunit
FT /note= "contains an intron"

```

```

FT intron 1259..1331
FT /tag= d
FT polyA_site 1797..2024
FT /tag= e
FT /note= "Bovine growth hormone polyA site"
FT promoter 2110..2737
FT /tag= f
FT /note= "CMV promoter"
FT iDNA 2738..2919
FT /tag= g
FT CDS 2983..3990
FT /tag= h
FT /product= p40 subunit
FT polyA_site 4075..4306
FT /tag= i
FT /note= "bovine growth hormone polyA site"
FN W09746263-A1.
PD 11-DEC-1997.
PF 04-JUN-1997; U09591.
PR 05-JUN-1996; US-659206.
PA (AURA) AUPAGEN INC.
PI Rakhmillevich AL Yang N;
DR WPI: 98-041898/04.
DR P-PSDB: W44004-05.
PT Interleukin-12 gene therapy of tumours - comprises delivering
PT construct containing promoter and sequences encoding interleukin-12
PT p35 and p40 subunit(s) to target cells in vivo
PT Claim 5; page 24-30; 50pp; English.
CC Plasmid pWRG3169 is a tandem plasmid encoding both subunits, i.e.
CC p35 (see W44004) and p40 (see W44005), of murine interleukin-12.
CC (IL-12). Each subunit gene was cloned from a mouse spleen cDNA
CC library and has been placed under the transcriptional control of
CC a separate cytomegalovirus (CMV) promoter. An SV40 splicing
CC donor/splicing acceptor is provided between each subunit gene and
CC its CMV promoter. The backbone of the plasmid is from pUC19.
CC pWRG3139 induces at least twice the expression of IL-12 as the
CC bicistronic vector pWRG3136 (see V02043) in vivo and in vitro. A
CC novel method of treating tumours in a mammal involves delivering
CC copies of an expressible foreign genetic construct, especially
CC pWRG3169 or pWRG3196, comprising a promoter operative in the
CC mammalian epidermal cells and DNA sequences encoding p35 and p40
CC subunits of IL-12 to target cells in vivo. Delivery of the
CC construct allows IL-12 expression for treatment of solid,
CC metastatic or disseminated tumours, and regression of established
CC tumours. The treatment is effective even when the genetic
CC construct is delivered to a site distant from the tumour.
CC Sequence 7287 BP; 1798 A; 1877 C; 1766 G; 1765 T;
SQ

```

```

Query Match 41.9%, Score 2141, DB 39, Length 7287;
Best Local Similarity 99.4%; Pred. NO 0.00e+00;
Matches 2155; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Db 5004 aggtggagcacttttcgggaaatgacgagaaacccctattttatttttctaaataca 5063
QY 2904 AGGTGGCACTTTTCGGGAAATGTGCGGGAACCCCTATTGTGTTTCTTAATATACA 2963
Db 5064 ttcaaatatgtatccgctcatgagacaataaacctgataaatgcttcaataattgaaa 5123
QY 2964 TTCAAATATGTATCCGTCATGAGACAATAACCTGTATAATGCTTCAATTAATGAAA 3023
Db 5124 aagggaagatgagatattcaacattccgtgctgcgccctattcccttttttgcgcatt 5183
QY 3024 AAGGAAGAGATATGATATCAACATTTCTGTGTGCGCTTATTCCTTTTTCGGGAT 3083
Db 5184 ttgccttcctgttttgcctccaccagaacacgtgtaaaataaaagatgctgaagatca 5243
QY 3084 TTGCTTCTCTGTTTGTCTACCCAGAAACGCTGTGTGAAAGTAAAGATGCTGAAGATCA 3143
Db 5244 gttgggtgacagtggtttacatcagactgagatctcacacacgagtaagatccttgagac 5303
QY 3144 GTTGGGTGACGAGTGGGTATCATATCGAACTGATCTCAACAGCGGTAAGATCTCTTGAAG 3203
Db 5304 ttttcgcccgaagaacgttttccaatgatagacacatttttaagttctgtatgtgagcgc 5363

```











Cp	4152	ATGTTTCTACAGGGTCTTCAGGCTCAGT36AACGAAAACTCACGTTAAGGGATTTTGGTC	3999
Db	8894	atdagattacaagaaggatcttcacctagatccctttcaaatataaatgaagttttaa	8953
Cp	3492	ATGAGATTATCAAAAAGGATCTTCACCTAGATCCCTTTTAAATTAAAAATGAAGTTTAAA	3933
Db	8954	tcaatctaaagtataataagaataaacttgctgacacgttaccacgttaacttaactgaag	9013
Cp	3832	TCAATCTAAAGTATATATAGTAAGAACTTGGCTGACAGTTACCAATGCTTAACTAGTGAG	3873
Db	9014	gcacctctccagcgatctgtctatttggttcaatccatagttgctgactcccgcgtcgtg	9073
Cp	3872	GCACCTATCTCAGCGATCTGTCTATTCTGTTCACTCCATAGTTGCTGACTCCCGCTGCTG	3813
Db	9074	taataaactgacatacgggaagggttaacatctggcccccagtgctgcaatgatacgcgga	9133
Cp	3812	TATATAAATACGATACGGAGGGCTTACATCTGGCCCACTGTGTGAATGATACCGGCA	3753
Db	9134	gacccacgctccacggctccagattatacagcaataaacacgacgacgcggaagggccgag	9193
Cp	3752	GACCCAGGCTCACGGCTCCAGATTATCAGCAATAAACACGACGCGGAGAGGCCGCGAG	3693
Db	9194	gagadaagtggtctgaaacttatccggctccatccaggtctattaattgttgcgggaa	9253
Cp	3692	CGAGAAATGCTCTGAATTTATCCGCTTCATCCAGTCTATTAATTTGTCGGGAA	3633
Db	9254	gctagaatgaagtgcgcacattaagtgttcacaagggttgctgccattgatacagcc	9313
Cp	3632	GCTAGATAAGTAAGTTTCCGCAGTAAGAAGTTTCCGCAAGTTGTGGCATTTGCTACAGGC	3573
Db	9314	atgtgtgtgcacgctgcgttggttgatagggttcatcagctccggttcccaacgataca	9373
Cp	3572	ATGCTGTGTCAAGTCTGCTGTTGGTATGCTTCACTCAGTCCGGTTCGCAACGATCA	3513
Db	9374	agcgcauttacatgatcccccatgttgttcacaaaagcggttagctccttggtctctccg	9433
Cp	3512	AGGCGAGTTACATGATCCCCCATTTGTCGCAAAAAGCGGTAGTCTCTCGTCTCCG	3453
Db	9434	atcgttctcgaaagttaagtgtggcgagtggttatcactcatggttatgacgaactcgat	9493
Cp	3452	ATGCTTCTCAAGTAAGTTGGCTGAGATGATGCTTCACTCATGTTATGGCAGCACTGCAT	3393
Db	9494	aattctctactgcatgccatccgttaagatcctttcttgactggttgactcaaac	9553
Cp	3392	AAATCTCTTACTCTATGTCATCCGTAAGATGCTTTCTTGACGTGGTAGTACTCAACC	3333
Db	9554	aadtcatctcgaaaatagtgatgagcgacgcagtgctcttgcgcgggctcaatacgg	9613
Cp	3332	AAATCATCTCTGACAAATACGCGGCTCGGCAACGAGTTGCTTGGCCGCGCTCAATACGG	3273
Db	9614	gataataccgcgcacatacagagaacttttaaagtgctcatcatcttgaaaaacgttctcg	9673
Cp	3272	GATAATATCTATGACATATAGAACTTTAAAGTCTCTCATCTTGGAAAAACGTTCTTCC	3213
Db	9674	gcgcgaataacttcaagatcttcacgcgtttgaatccatttgatgataaccactcgtt	9733
Cp	3212	GCCTGAAAAACTCTCTAAGATATTCTATGCTGTGACATCCAGTTGATTAACCCACCTCGT	3153
Db	9734	gcacacacgtgatcttcagcatctttacttccacgcggtttctgggtgagcaaaaaca	9793
Cp	3152	GCACCCAGCTCATCTTCAGCATCTTTACTTTCACCCAGCGTTTCTGGGTGAGCAAAAACA	3093
Db	9794	ggaaggcaaaatgcgcgaaaaaaagggaataaggcgacacacgggaatgttgataactcata	9853
Cp	3092	GGAAGCTAAAATGCGGCAAAAAAGGAAATGAAGGGCGSAACGGAAATGTTGAATACTCAIA	3033
Db	9854	ctctctcttttttaattattgaagcatcttatcaaggttatgtctcatgagcgagatc	9913
Cp	3032	CCTCTCTTTTTCAAATATATTGACGCAATTTATCAGGGTTATTGTCTCATGAGCGGATAC	2973
Db	9914	ataattgaatgattttagaaaaataacaaaatagaggttcctccgcacattctcccgaaaa	9973
Cp	2972	ATATTCTAATGTATTATGAAAAATAACCAAAATAGAGGTTCTCGCGGCACATTTCCCGGAAA	2913

```

Db  9974  gtgccacctg  9983
      |||||
Cp  2912  GTGCCACCTG  2903

RESULT  14
ID  Q40279 standard: DNA: 4145 BP.
AC  Q40279;
DT  02-AUG-1993 (first entry)
DE  Sequence of clone pS2gpt-S4.
KW  Plasmid; cloning; restriction site; ss.
OS  Synthetic.
FH  Key      Location/Qualifiers
FT  misc_feature 1..2226      /tag= a
FT  /label= pN2gpt-S4
FT  /note= "Posn. 1 corresp. to the first nucleotide
FT  G '5-TGCACACTT TCGGGGAAT-3'."
FT  misc_feature 2227..2236
FT  /tag= b
FT  /label= SmaI adaptor
FT  /note= 2396..22851
FT  cds          2396..22851
FT  /tag= c
FT  /label= E. coli gpt gene
FT  /note= complement (3081..3323)
FT  promoter     3358..3451
FT  /tag= d
FT  /label= vaccinia p7.5 promoter
FT  /tag= e
FT  /label= S4 of pN2gpt-S4
FT  /note= "oligo p-atp(9)"
FT  misc_feature 2237..4145
FT  /tag= f
FT  /label= pN2gpt-S4
FT  AU9221269-A.
PD  04-MAR-1993.
PF  25-AUG-1992; G31269.
PR  26-AUG-1991; US-750080.
PR  20-JUL-1992; US-914739.
PI  (IMMO ) IMMUNO AG.
PI  Dörner F., Falkner FG., Pfeleiderer M., Schelflinger F.;
DR  WPI; 93-126451/16.
PT  Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PT  direct molecular cloning of modified DNA molecule contg.
PI  cytoplasmic DNA virus genome
PS  Example; Pages 160-162; 206pp; English.
CC  Plasmids pN2gpt-S3A and pN2gpt-S4 comprise expression cassettes
CC  with a selective marker. These plasmids were constructed by first
CC  making plasmids pN2-gptA and pN2-gptB which contain an E. coli gpt
CC  gene driven by the vaccinia virus P7.5 promoter, flanked by
CC  several unique restriction sites including NotI. The plasmid
CC  pN2gpt-S4 was digested with XbaI and ligated with a SmaI-adaptor
CC  inactivating the XbaI and creating a SmaI site. The resulting
CC  plasmid was designated pS2gpt-S4. The S4-promoter segment of
CC  lasmid pS2gpt-S4 was removed by cleavage with PstI and HpaI and
CC  replaced with a 172 bp PstI-HpaI P2-promoter segment. The final
CC  plasmid was designated pS2gpt-P2.
CC  Sequence 4145 BP.
SQ  Sequence 4145 BP. 1067 A. 1007 C. 1039 G. 1032 T.

Query Match 41.8%; Score 2138; DB 7; Length 4145.
Best Local Similarity 99.4%; Pred. No 0 00e+00;
Matches 2152. Conservative 0; Mismatches 14; Indels 0; Gars

```

4026 GAAAGATATAGTATTCAGATATTTGGTGTGGGATATATCTTTTGGGGCATTTT 3085  
 DB |  
 4131 GATTTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 240  
 DB |  
 4086 GGTCTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 3145  
 QY |  
 DB |  
 241 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 309  
 DB |  
 4146 TGGTGGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3205  
 QY |  
 DB |  
 4301 TGGTGGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
 DB |  
 4296 TGGTGGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3265  
 QY |  
 DB |  
 461 TATATCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 420  
 DB |  
 4266 TATATCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 3325  
 QY |  
 DB |  
 421 ATGCTTGGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
 DB |  
 4026 ATGCTTGGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3385  
 QY |  
 DB |  
 481 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
 DB |  
 4486 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3445  
 QY |  
 DB |  
 541 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
 DB |  
 4446 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3505  
 QY |  
 DB |  
 601 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
 DB |  
 4506 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3565  
 QY |  
 DB |  
 661 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
 DB |  
 4566 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3625  
 QY |  
 DB |  
 721 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
 DB |  
 4626 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3685  
 QY |  
 DB |  
 781 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
 DB |  
 4686 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3745  
 QY |  
 DB |  
 841 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
 DB |  
 4746 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3805  
 QY |  
 DB |  
 901 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 DB |  
 4806 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3865  
 QY |  
 DB |  
 961 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
 DB |  
 4856 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3925  
 QY |  
 DB |  
 1021 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
 DB |  
 4926 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3985  
 QY |  
 DB |  
 1081 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
 DB |  
 4986 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4045  
 QY |  
 DB |  
 1141 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
 DB |  
 4046 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4105  
 QY |  
 DB |  
 1201 GAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
 DB |

4106 CAATAAACACACGCTACACGCTGATGATGATGATGATGATGATGATGATGATGAT 4165  
 DB |  
 1261 TGGTGGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 DB |  
 4166 TGGTGGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4225  
 QY |  
 DB |  
 1321 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
 DB |  
 4226 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4285  
 QY |  
 DB |  
 1381 TGGTGGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
 DB |  
 4286 TGGTGGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4345  
 QY |  
 DB |  
 1441 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
 DB |  
 4346 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4405  
 QY |  
 DB |  
 1501 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
 DB |  
 4406 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4465  
 QY |  
 DB |  
 1561 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
 DB |  
 4466 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4525  
 QY |  
 DB |  
 1621 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
 DB |  
 4526 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4585  
 QY |  
 DB |  
 1681 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
 DB |  
 4586 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4645  
 QY |  
 DB |  
 1741 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1800  
 DB |  
 4646 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4705  
 QY |  
 DB |  
 1801 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1860  
 DB |  
 4706 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4765  
 QY |  
 DB |  
 1861 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
 DB |  
 4766 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4825  
 QY |  
 DB |  
 1921 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1980  
 DB |  
 4826 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4885  
 QY |  
 DB |  
 1981 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2040  
 DB |  
 4886 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4945  
 QY |  
 DB |  
 2041 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2100  
 DB |  
 4946 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2005  
 QY |  
 DB |  
 2101 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2160  
 DB |  
 5006 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2220  
 QY |  
 DB |  
 2161 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2280  
 DB |  
 5066 GATTTCTCTCTTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2340  
 QY |

RESULT 15

ID T69891 standard; DNA: 4540 BP.

AC T69891;

DT 19-SEP-1997 (first entry)

DE Homo. cyclin D1; human. 5964 bp. 1436 bp. 1436 bp. 1436 bp.

KW Cyclin D1; cyclin-dependent protein kinase 4; CDK4; cell cycle; etc.

OS Chimaeric Homo sapiens:  
 FH Chimaeric synthetic:  
 Key 2218..4335  
 CDS /tag- a  
 PN W09725345.A1.  
 PD 17-JUL-1997.  
 PF 03-JAN-1997: U00140.  
 PR 05-JAN-1996: US-009629.  
 PA (ELIL J LILLY & CO ELI.  
 PI Kovacevic S, Otto KA, Rao RN:  
 DR WPI: 97-372814/34.  
 DR P-PSDB: W18573.  
 PI Fusion protein comprising human cyclin and cyclin-dependent protein  
 PT kinase - useful for research in cell cycle regulation allowing  
 PT addition of both components simultaneously, giving greater control  
 PT over reaction conditions  
 PS Disclosure: Page 17-22: 57pp. English.  
 CC A DNA sequence (J69891) in plasmid K480, deposited in E. coli as  
 CC NPPI R-21491, includes a coding region for a fusion protein  
 CC (W18573) comprising human cyclin D1 linked to human CDK4, flanked  
 CC by N-terminal histidine residues, a myc epitope and a thrombin  
 CC cleavage site, and by a C-terminal streptavidin binding domain, to  
 CC facilitate purification. The fusion protein can be expressed in  
 CC host cells, pref. using a baculovirus expression system, for use in  
 CC research into cell cycle regulation.  
 SQ Sequence 4540 BP: 1067 A: 1253 G: 1040 T:  
 Query Match 41.8% Score 2138: DB 32: Length 4540:  
 Best Local Similarity 99.4%: Prod No 0 00e-00:  
 Matches 2154: Conservative 0: Mismatches 16: Indels 0: Gaps 0:  
 Db 70 cagtgacacttttcggggaatgtcgcgggaacccctatttggttatttttctaaataac 129  
 Qy 2903 CAGSISGACATTTTGGGGAATGTCGCGGAACTTGTGTAATTTTTCATTAATAC 2962  
 Db 130 attcaaatatgcctcgtatgaacataaacctcgtatgaatgcttcaataattgaa 189  
 Qy 2963 ATTCAAAATGTAATGCTTATGAAATATGTAATGTAATGTAATGTAATGTAAT 3022  
 Db 190 aaagaaagatgagttatcaacatttcggtgctgcgccttattccctttttgggcat 249  
 Qy 3023 AAAGTAACATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 3082  
 Db 250 ttgccttcctgttttgcacccagaaacgctggtgaaatgaaagatgctgaagtc 309  
 Qy 3083 TTGCTTCTGCTTTTGTGCTGCTTGTGCTGCTTGTGCTGCTTGTGCTGCTTGTGCTG 3142  
 Db 310 agttggatgcaagtggtttacatgaactgagatcgaacagcggtgaagatccttgaga 369  
 Qy 3143 AGTTGGTGTGACAGTGGGTTACATGAACTGCAACAGCGGTGAAGATCCTTGAGA 3202  
 Db 370 gtttcgcgcgcgaagcgttttccaatgataggaactttttaaagtctcgtatggcg 429  
 Qy 3203 GTTTTCGCGCGGGAAGACGTTTTCGAATGATGAGCACTTTTAAAGTCTGCTATGCTA 3262  
 Db 430 cgttattatccctattgacgcggggaagaaacactcgttcgcgcgcatacactatttc 489  
 Qy 3263 CACTATTATTCGCTATTGAGCGCGGGAAGCAAGCAACTGCGGTGCGCGCGGTATTCTC 3322  
 Db 490 aqaatgacttggttgatctaccagtcacagaaagacatcttacggatggcatgaacag 549  
 Qy 3323 ATAAATGACTTGTGTGATATGATATGATATGATATGATATGATATGATATGATATG 3382  
 Db 550 taagaaattatgcagtgctgcataaccatgagtgatgaacactgcggcgaacttacttc 609  
 Qy 3383 TAAAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3442  
 Db 510 tcaacagatcgaagaccgaagagactaacgccttttttgcacacatcgggggatcatg 669  
 Qy 3443 TGACAAAGATCGGAGGAGGCAAGGAGGCAAGGAGGCAAGGAGGCAAGGAGGCAAGG 3502  
 Db 570 taactcgccttgatcgttgggaacccggagctgaatgaagaaacccaataccaaagcggagtg 729

Qy 3503 TAACTCGCTTATGCTTGGGAACCGGAGCTTCAATGAAAGCAATCAAAACAAACGAGGCTG 3562  
 Db 730 acacaacgatgctgtatagaatggcaacaacgcttgcgcaaacatttactcgcgaactac 789  
 Qy 3563 ACACACGSAIGCTTGAACAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 3622  
 Db 790 ttacttaacttcccgccgaacaataatagactgagatgagagcgagataaagcttcagagac 849  
 Qy 3623 TTACTCTAGCTTCCCGGGAATTAATTAATAGACTGGAATGGAATGGAATGGAATGGA 3682  
 Db 850 cacttctgctgcgcgccttcgcgcgccttgcgcgccttgcgcgccttgcgcgccttgcgc 909  
 Qy 3683 CACTTCTGAGCTGCTGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3742  
 Db 910 agcgtgagctgcgcgccttgcgcgccttgcgcgccttgcgcgccttgcgcgccttgcgc 969  
 Qy 3743 AGCGTGGGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3802  
 Db 970 tagttatctacacgagcggggagctgcgcgccttgcgcgccttgcgcgccttgcgcgcct 1029  
 Qy 3803 TAGTTATCTACAGACGCGGGAGTCAAGCACTATGGATGATGATGATGATGATGATGAT 3862  
 Db 1030 agataggtgcctcactgattgaagcatttgactgcacacacacacacacacacacacac 1089  
 Qy 3863 AGATAGTGCCTCACTGATTAAAGCATTTGTAAGTCTGTAAGTCTGTAAGTCTGTAAGT 3922  
 Db 1090 tttaagatttaataaacttcttcttcttcttcttcttcttcttcttcttcttcttctt 1149  
 Qy 3923 TTAGATGATTTAAACCTTCAATTTTAAATTAAGGATCTAGGTAAGGATCTAGGTA 3982  
 Db 1150 ataactcctgacccaaatcccttaacgtgagtttcttccactgagcgtcagaccccg 1209  
 Qy 3983 ATAATCTCATGACCAAAATCCCTTAACGTGAGTTTCCTTCCACTGAGCGTCAAGCCCG 4042  
 Db 1210 tagaaagatcaaaagatcttcttgagatccttttttctgcgcgccttgcgcgccttgcgc 1269  
 Qy 4043 TAGAAAGATCAAAAGATCTTCTTGAATCTCTTTTCTGCGGTAATGCTGCTGCTGCTG 4102  
 Db 1270 aaacaaaaaaccccgctaccagcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1329  
 Qy 4103 AAACAAAAAACCCGCTACCGAGCGGTGTTGTTGCGGATCAAGAGTACCAACTC 4162  
 Db 1330 ttttccgaggttaactggcttcagcagcagcagcagcagcagcagcagcagcagcagcag 1389  
 Qy 4163 TTTTCCGAAGGTAACTGCTTTCAGCAGAGCGCAGATACCAAAATCTGCTGCTTCTAGTGT 4222  
 Db 1390 agcgtgagtgagcgcacacttcaagaactctgttagcgcgcgcctacatacctcgccttgc 1449  
 Qy 4223 AGCGTAGTTAGCGCACTTCAAGAACTCTGTATGTAAGTCTGTAAGTCTGTAAGTCTG 4282  
 Db 1450 taactcgtttaccagtgctgcgcagtgccagtgccagtgccagtgccagtgccagtgccag 1509  
 Qy 4283 TAATCTGTATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4342  
 Db 1510 caagcagtagttaccggataaggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1569  
 Qy 4343 CAAGAGATAGTTACCGGATAAGGCGGCAAGCGCTGCGCTTGAAGCGGCGCTGCGCTG 4402  
 Db 1570 agccagcttgagcgaacacactacacgcgaactgaatacctacagcgtgaagcattgag 1629  
 Qy 4403 AGCGGCTGTTGAGCGGAAGCACTTCAAGAACTCTGTATGTAAGTCTGTAAGTCTG 4482  
 Db 1630 aaagcgcgcgccttcccgaaaggagaaaggcgagcagctatcccgtaacgcgcgcgcgcgc 1689  
 Qy 4463 AAAGCGCGCGCTTCCCGAAGCGGCAAGCGGCAAGCGGCAAGCGGCAAGCGGCAAG 4522  
 Db 1690 gaacagcagc 1749  
 Qy 4523 GAACAGAGAGCGGCAAGCGGAGCTTCCAGGCGGCAAGCGGCAAGCGGCAAGCGGCAAG 4582  
 Db 1750 tccggtttccgcacacttgcacttgagcgtcgtatttttgcgcgcgcgcgcgcgcgcgcgcgc 1809

QY 4583 TGGGTTTGGGCACTCTGCACTTACAGCTGTGATTTTTGATGATGCTCTGCAAGGAGGAGGAGGGA 4542  
Db 1810 ggcctatgnaaaagcccaagcaagcgccttttttaagcttccctagccttttttgcctggcctt 1869  
TT  
QY 4643 GGTATATGAAAAAGCTGATGAGGAGGCTTTTTTAAGGCTTCTGAGGCTTTTTGCTGGGCTT 4702  
Db 1870 ttgctcacatgttcttctcgtttatccctctgattctgtatgataccgctattaccgcct 1929  
TT  
QY 4704 TTTCTACAGTGTCTTTCTGCTGTTATGCTCTGATTCTGTGATAACCGCTATTACCGCCT 4762  
Db 1930 ttaagtgaatctatagcgtctgctgagagcaagcaagcaagcaagcaagcaagcaagcaagca 1989  
TT  
QY 4763 TTTGATGATCTGATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4822  
Db 1990 aggaagcgaag 2049  
TT  
QY 4823 AGCAAGAGCAAG 4882  
Db 2050 gatgagagctggag 2109  
TT  
QY 4883 AATGAGCTTGGATGATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4942  
Db 2110 atagtgattagctcct 2169  
TT  
QY 4943 ATGCTAGTTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5002  
Db 2170 tcttctctcgaattctcgaagcgaatgacaaatttcacacaggaagaaacagctatgaccatg 2229  
TT  
QY 5004 TTTTCTGTGAATTTCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 5062  
Db 2230 agcccaagct 2239  
TTTTTTTTTTTT  
QY 5063 AGGCCAAGCT 5072

Search completed: Mon Mar 15 17:28:59 1999  
Job time: 1082 secs.

\*\*\*\*\*

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPS Research n.d. - n.a. database search, using Smith-Waterman algorithm  
Run on: Mon Mar 15 14:33:11 1999; MasPar time 6493.81 Seconds  
Tabular output not generated. 1410.547 Million cell updates/sec

Title: >US-09-020-716-5  
Description: (1-5115) from US09020716.seq  
Perfect Score: 5115  
N.A. Sequence: 1 GTTGGAGGCTCTCCCATATG .....TCAAGCTATGATCCCAACGC 5115  
Comp: CAACCTCTGAGAGGGGTATAC .....AGTTGATAGCTAGCTTGGC

Scoring table: TABLE default  
Gap 5

Nmatch STD : Dbase 0: Query 0

Searched: 2275926 seqs, 895388244 bases x 2

Post-processing: Minimum Match 08

Listing first 45 summaries

Database: emb1-est56  
Database: 1:em\_est1 2:cm\_gss1 3:em\_gss2 4:em\_gss3  
genbank-est109

5:gb\_est1 6:gb\_est10 7:gb\_est11 8:gb\_est12 9:gb\_est13  
10:gb\_est14 11:gb\_est15 12:gb\_est16 13:gb\_est17  
14:gb\_est18 15:gb\_est19 16:gb\_est2 17:gb\_est20  
18:gb\_est21 19:gb\_est3 20:gb\_est4 21:gb\_est5 22:gb\_est6  
23:gb\_est7 24:gb\_est8 25:gb\_est9 26:gb\_gss1 27:gb\_gss2  
28:gb\_gss3 29:gb\_gss4

Statistics: Mean 12.493; Variance 1.996; scale 6.260

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	555	12.8	721	26	AG004662 Homo sapiens genomic D	0.00e+00
2	555	12.8	721	27	AG011001 Homo sapiens genomic D	0.00e+00
3	538	12.5	711	27	AG010947 Homo sapiens genomic D	0.00e+00
4	538	12.5	711	26	AG004508 Homo sapiens genomic D	0.00e+00
5	533	12.4	718	26	AG004363 Homo sapiens genomic D	0.00e+00
6	533	12.4	718	27	AG010489 Homo sapiens genomic D	0.00e+00
7	530	12.3	692	27	AG040352 CIT-HSP-2327K21.TF CIT	0.00e+00
8	528	12.3	698	27	AG009976 Homo sapiens genomic D	0.00e+00
9	528	12.3	698	26	AG003787 Homo sapiens genomic D	0.00e+00
10	624	12.2	720	28	AG013858 Homo sapiens genomic D	0.00e+00
11	624	12.2	720	26	AG006062 Homo sapiens genomic D	0.00e+00
12	517	12.1	652	28	AG018633 CIT-HSP-2379M3.TF CIT	0.00e+00
13	519	12.1	747	25	AG007052 Homo sapiens genomic D	0.00e+00

C	14	612	12.0	666	27	AQ079096	CIT-HSP-2356B24.TF CIT	0.00e+00
C	15	609	11.9	540	27	AQ074298	CIT-HSP-2382L21.TF CIT	0.00e+00
C	16	598	11.7	596	27	AG009765	Homo sapiens genomic D	0.00e+00
C	17	596	11.7	596	26	AG003576	Homo sapiens genomic D	0.00e+00
C	18	596	11.7	748	17	A1124281	1 53 thrsophila 8-12 h	0.00e+00
C	19	591	11.6	542	28	AQ13212	CIT-HSP-2372L22.TF CIT	0.00e+00
C	20	592	11.6	690	27	AG009464	Homo sapiens genomic D	0.00e+00
C	21	595	11.6	703	26	AG001761	Homo sapiens genomic D	0.00e+00
C	22	581	11.4	594	26	C1786	Ciona intestinalis gen	0.00e+00
C	23	579	11.3	634	27	AQ38010	CIT-HSP-2328L2.TF CIT	0.00e+00
C	24	578	11.3	653	27	AQ37504	CIT-HSP-2365K21.TF CIT	0.00e+00
C	25	577	11.3	635	26	AG002183	Homo sapiens genomic D	0.00e+00
C	26	573	11.2	633	27	AQ04787	CIT-HSP-2336K22.TF CIT	0.00e+00
C	27	574	11.2	661	27	AQ057603	CIT-HSP-2345M3.TF CIT	0.00e+00
C	28	569	11.1	637	27	AQ077360	CIT-HSP-2365E1.TF CIT	0.00e+00
C	29	564	11.0	615	28	AQ112424	CIT-HSP-2376F3.TF CIT	0.00e+00
C	30	563	11.0	625	28	AQ111342	CIT-HSP-2371M3.TF CIT	0.00e+00
C	31	559	10.9	689	28	AG014394	Homo sapiens genomic D	0.00e+00
C	32	547	10.7	600	28	AQ111411	CIT-HSP-2373F2.TF CIT	0.00e+00
C	33	545	10.7	623	27	AQ081172	CIT-HSP-2357C31.TF CIT	0.00e+00
C	34	540	10.6	612	27	AQ080586	CIT-HSP-2382D31.TF CIT	0.00e+00
C	35	530	10.4	605	27	AQ077353	CIT-HSP-2365C21.TF CIT	0.00e+00
C	36	533	10.4	682	27	AQ074593	CIT-HSP-2361L23.TF CIT	0.00e+00
C	37	526	10.3	587	27	AQ079166	CIT-HSP-2382M2.TF CIT	0.00e+00
C	38	526	10.3	594	26	AG007826	Homo sapiens genomic D	0.00e+00
C	39	525	10.3	604	26	AG002706	Homo sapiens genomic D	0.00e+00
C	40	525	10.3	629	27	AG011367	Homo sapiens genomic D	0.00e+00
C	41	528	10.3	743	26	AG002826	Homo sapiens genomic D	0.00e+00
C	42	524	10.2	540	27	AQ074307	CIT-HSP-2382N1.TF CIT	0.00e+00
C	43	521	10.2	597	27	AQ040643	CIT-HSP-2324K24.TF CIT	0.00e+00
C	44	510	10.0	664	26	AG005286	Homo sapiens genomic D	0.00e+00
C	45	510	10.0	664	27	AG012931	Homo sapiens genomic D	0.00e+00

## ALIGNMENTS

RESULT	LOCUS	DEFINITION	721 bp	DNA	GSS	30-JAN-1998
1	AG004662	Homo sapiens genomic DNA, 21q region, clone: 879G12X91, genomic survey sequence.				
	AG004662					
	g2826191					
	GSS					
	Homo sapiens					
	Homo sapiens DNA, clone: 879G12X91.					
	KEYWORDS					
	SOURCE					
	ORGANISM					
	REFERENCE					
	AUTHORS					
	TITLE					
	JOURNAL					
	REFERENCE					
	AUTHORS					
	TITLE					
	JOURNAL					
	BASE COUNT		195 a	163 c	179 g	179 t
	ORIGIN					9 others
	FEATURES					
	Source					
	Location/Qualifiers					
	1..721					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	/chromosome="21"					
	/clone="879G12X91"					
	/map="21q"					
	Query Match					12.8%, Score 659, DB 26, Length 721;
	Best Local Similarity					94.6%; Pred No 0.00e+00;
	Matches					582, Conservative 0, Mismatches 6, indels 4, gaps 4;

Submitted (30-JAN-1998) to the DDBJ/EMBL/GenBank databases.  
Masahira Hattori, Kitasato University, Department of Science, JST  
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan  
(E-mail: hattori@bc.ims.u-tokyo.ac.jp. Tel: 0427-78-9732.  
Fax: 0427-78-9561)





[illegible]







Db	570	CGAAGTATAC	-TAIAIACTTTAGATTGATTTAAACTTCATTTTAAATTAAGGAT	628			
Qy	3903	CGAAGTTTACTCATATATACTTTAGATTGATTTAAACTTCATTTTAAATTAAGGAT	3962				
Db	629	CTAGGTGAAGATGCTNTTTT-GATATGCTCATGACCAAAATCCCTTAAGCTGAGTTTCGTT	687				
Qy	3963	CTAGGTGAAGATGCTTTTGTATATCTCATGACCAAAATCCCTTAAGCTGAGTTTCGTT	4022				
Db	688	TCACATGA	695				
Qy	4023	-CCACTGA	4029				
RESULT 10							
LOCUS	AG013858	720 bp	DNA	GSS 10-SEP-1998			
DEFINITION	Homo sapiens genomic DNA, 21q region, clone: 762015SpN14, genomic survey sequence.						
ACCESSION	AG013858						
NID	q3560348						
KEYWORDS	GSS						
SOURCE	Homo sapiens						
ORGANISM	Homo sapiens DNA, clone:762015SpN14.						
REFERENCE	Eukaryota: Metazoa: Chordata: Vertebrata: Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.						
AUTHORS	1 (bases 1 to 720)						
TITLE	Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.						
JOURNAL	Homo sapiens genomic DNA, chromosome 21q						
REFERENCE	Published Only in DataBase (1998) In press						
AUTHORS	2 (bases 1 to 720)						
TITLE	Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.						
JOURNAL	Submitted (10-SEP-1998) to the DDBJ/EMBL/Genbank databases. Masahira Hattori, Kitasato University, Department of Science, Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan (E-mail:hattori@hpq.ims.u-tokyo.ac.jp. Tel:0427-78-9732, Fax:0427-78-9561)						
FEATURES							
source							
1 720							
/organism="Homo sapiens"							
/db_xref="taxon:9606"							
/chromosome="21"							
/clone="762015SpN14"							
/map="21q"							
BASE COUNT	172 a	178 c	170 g	186 t			
ORIGIN							
Query Match	10.2%, Score 624, DB 28, Length 720;						
Best Local Similarity	97.2%, Fred No. 3, qje-03;						
Matches	571: Conservative	0: Mismatches	13: Indels	6: Gaps			
Db	35	AGCTGCTCGTGGCTCTCTCTGTTTCGAGACCTTCGCGCTTACCGGATACCTGTCGCGCTTT	94				
Cp	4547	AGCTGCTCTGTCGCTCTCTCTGTTTCGAGACCTTCGCGCTTACCGGATACCTGTCGCGCTTT	4488				
Db	95	CTGCTTTCGGGAAGGTGGCGCTTCTCATGCTACGCTGAGTATCTCAGTTCGGTG	154				
Cp	4487	CTGCTTTCGGGAAGGTGGCGCTTCTCATGCTACGCTGAGTATCTCAGTTCGGTG	4428				
Db	155	TAGTTCGTTTCGCTCCAACTGGGTGTGTGTGACGAACCCCGCTTCACCGCCGACCGGTGC	214				
Cp	4427	TAGTTCGTTTCGCTCCAACTGGGTGTGTGTGACGAACCCCGCTTCACCGCCGACCGGTGC	4368				
Db	215	GCGTTATCGGTAACTATCGCTTGTAGTCCACCCCGGTGAAGACAGCACTTATCCCACTG	274				
Cp	4367	GCGTTATCGGTAACTATCGCTTGTAGTCCACCCCGGTGAAGACAGCACTTATCCCACTG	4308				
Db	275	GCAGCAGCCACTGGTAAACAGATTAAGCAGACGAGGTATGTAGCGGTGCTACAGAGTTC	334				
Cp	4307	GCAGCAGCCACTGGTAAACAGATTAAGCAGACGAGGTATGTAGCGGTGCTACAGAGTTC	4248				
Db	335	TTGAATGGTGGCCTAAATAAGGTAACATAGAAAGAACAGATTTGGTATCTCGGCTCTG	394				



**TITLE** Homo sapiens genomic DNA, chromosome 21q  
**JOURNAL** Published only in Database (1998) In press  
**REFERENCE** 2 (bases 1 to 747)  
**AUTHORS** Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (17-APR-1998) to the DDBJ/EMBL/GenBank databases.  
 Masahira Hattori, Kitasato University, Department of Science, JST  
 Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan  
 (E-mail: hattori@hgc.ims.u-tokyo.ac.jp, Tel: 0427-78-9732,  
 Fax: 0427-78-9551)  
**FEATURES** Location/Qualifiers  
 source  
 1..747  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="21"  
 /clone="c48B2A3"  
 /map="21q"  
**BASE COUNT** 194 a 171 c 176 g 194 t 12 others  
**ORIGIN**  
 Query Match 12.1%; Score 619; DB 26; Length 747;  
 Best Local Similarity 96.3%; Pred. No. 0.00e+00;  
 Matches 683; Conservative 0; Mismatches 17; Indels 9; Gaps 9;  
 Db 41 ATCTGGCTCTGCTGAAGCCAGTTACCTTCGGGAAAGAGTTGGTAGCTTTGATCCGGC 100  
 Cp 4199 ATCTGGCTCTGCTGAAGCCAGTTACCTTCGGGAAAGAGTTGGTAGCTTTGATCCGGC 4140  
 Db 101 AAACAACACCGCTGTAGTCTGIG-TTTTITGTTTAAACACACAGTAACGGCAGCA 159  
 Cp 4139 AAACAACACCGCTGTAGTCTGIG-TTTTITGTTTAAACACACAGTAACGGCAGCA 4080  
 Db 160 AAAAAAGGATCTCAAGAAGATCTTTTATCTTTTACGGGCTCTGAGCTCAGTGGAAAC 219  
 Cp 4079 AAAAAAGGATCTCAAGAAGATCTTTTATCTTTTACGGGCTCTGAGCTCAGTGGAAAC 4020  
 Db 220 GAAACATCAGTTAAGGATTTTGGTCATGAGATTAATCAAAAGATCTTACCTAGATC 279  
 Cp 4019 GAAACATCAGTTAAGGATTTTGGTCATGAGATTAATCAAAAGATCTTACCTAGATC 3960  
 Db 280 CTTTAAATTAATAAGTAACTTTTAAATCAATCTAAAGTATATATGAGTAAACCTTGGTCT 339  
 Cp 3959 CTTTAAATTAATAAGTAACTTTTAAATCAATCTAAAGTATATATGAGTAAACCTTGGTCT 3900  
 Db 340 GACAGTTACCAATGCTTAATCAGTAGAGGACCTATCTCAGCGATCTGTCTATTTCGTCA 399  
 Cp 3409 GACAGTTACCAATGCTTAATCAGTAGAGGACCTATCTCAGCGATCTGTCTATTTCGTCA 3840  
 Db 400 TCCATAGTTGCTGACTCCCGGCTGGTGTAGATAAATCAATACATACGAGGGGT-ACCATCT 458  
 Cp 3839 TCCATAGTTGCTGACTCCCGGCTGGTGTAGATAAATCAATACATACGAGGGGTATCCATCT 3780  
 Db 459 GGGGCGGAGTCTGCAATATACCGGAGAGGACAGCTACGCTACCGGCTCCAGATTATCAGCA 518  
 Cp 3779 GGGGCGGAGTCTGCAATATACCGGAGAGGACAGCTACGCTACCGGCTCCAGATTATCAGCA 3720  
 Db 519 ATAACACGAGCGCGGAGGGGCGGAGCGCA-AGCTGGTCTCTGCAACTTTATCGGCTC 577  
 Cp 3719 ATAACACGAGCGCGGAGGGGCGGAGCGCA-AGCTGGTCTCTGCAACTTTATCGGCTC 3661  
 Db 578 CATCCAGCTATTAAATGTTGGCGGGAAGCTAGAGTAAAGTAGTTTCGGCAGTTAATAGT 637  
 Cp 3660 CATCCAGCTATTAAATGTTGGCGGGAAGCTAGAGTAA-ATAGTT-CGGCAGTTAATAGT 3603  
 Db 638 TTTGGCAACGTTTGTGATTTGATAGAGG-ATGCTGGTGTGACGCTCGCTGCTTGGTA-T 3544  
 Cp 3602 TTTGGCAACGTTTGTGATTTGATAGAGG-ATGCTGGTGTGACGCTCGCTGCTTGGTA-T 3544  
 Db 638 GCTTCATTTTCASNTGCGGINTGCCAACATCAATCAAGGCGGTTAACATGAT 746  
 Cp 3543 GCTTCATTT-CAGTCTCGG-TTCCCAACGATCAAGGCGGATACATGAT 3497

**RESULT** 14  
**LOCUS** AQ079096 666 bp DNA GSS 19-AUG-1998  
**DEFINITION** CIT-HSP-2356E24.TF CIT-HSP Homo sapiens genomic clone 2356E24, genomic survey sequence.  
**ACCESSION** AQ079096  
**NID** g3440280  
**KEYWORDS** GSS.  
**SOURCE** human.  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
 Primates; Catarrhini; Hominoidea; Homo.  
**REFERENCE** 1 (bases 1 to 666)  
**AUTHORS** Adams, M.D., Pounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, M. and Venter, J.C.  
**TITLE** Use of a random human BAC End Sequence Database for Sequence-Ready Map Building  
**JOURNAL** Unpublished (1998)  
**COMMENT**  
 Contact: Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: mdadams@tigr.org  
 Clones are available from Research Genetics (info@resgen.com). BAC  
 end search page:  
 http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html.  
 Seq primer: M13-21  
 Class: BAC ends.  
**FEATURES** Location/Qualifiers  
 source  
 1..666  
 /organism="Homo sapiens"  
 /note="Vector: pBelBAC11; Site\_1: HindIII; Site\_2: HindIII"  
 /db\_xref="taxon:9606"  
 /clone="2356E24"  
 /clone\_lib="CIT-HSP"  
 /sex="Male"  
 /cell\_type="Sperm"  
**BASE COUNT** 156 a 186 c 174 g 149 t 1 others  
**ORIGIN**  
 Query Match 12.0%; Score 612; DB 27; Length 666;  
 Best Local Similarity 99.7%; Pred. No. 0.00e+00;  
 Matches 613; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Db 52 AGCTTGAGTATTCTATAGTGTGACCTTAATAGCTTGGGGTAAATCATGTGATAGCTGTTT 111  
 Cp 5102 AGCTTGAGTATTCTATAGTGTGACCTTAATAGCTTGGGGTAAATCATGTGATAGCTGTTT 5043  
 Db 112 CTTCTGTGAAATTTTATCGCTGACATTCACAACTACAACTACAACTACAACTACAACTACAACT 171  
 Cp 5042 CTTCTGTGAAATTTTATCGCTGACATTCACAACTACAACTACAACTACAACTACAACTACAACT 4983  
 Db 172 TGTAAAGGCTTGGGGTGGCTTAATGAGTGTGAGCTTAACATCATTAATTTGGCTTGGCTGACTG 231  
 Cp 4982 TGTAAAGGCTTGGGGTGGCTTAATGAGTGTGAGCTTAACATCATTAATTTGGCTTGGCTGACTG 4923  
 Db 232 CCGGCTTTCCAGTCCGGAACCTGTGGTCCGAGTATTAATGAGTGTGAGCTTAACATCATTAATTTGGCT 291  
 Cp 4922 CCGGCTTTCCAGTCCGGAACCTGTGGTCCGAGTATTAATGAGTGTGAGCTTAACATCATTAATTTGGCT 4863  
 Db 292 GGGAGAGGGGTTTGGCTATTGGGGGCTTTTGGGTTTGGGTTTGGGTTTGGGTTTGGGTTTGGGTTT 351  
 Cp 4862 GGGAGAGGGGTTTGGCTATTGGGGGCTTTTGGGTTTGGGTTTGGGTTTGGGTTTGGGTTTGGGTTT 4803  
 Db 352 TCGTCTGTTTGGGCTGTGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 411  
 Cp 4802 TCGTCTGTTTGGGCTGTGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4743  
 Db 412 ACAGATCAGGGGATTAACGCGAGGAAGACATGTGAGCAAAAGGCGCAAGAGGCGCAAGG 471





\*\*\*\*\*  
 M O E R E A  
 \*\*\*\*\*  
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution Rights by Oxford Molecular Ltd

MParch\_en n.a. - n a database search, using Smith-Waterman algorithm  
 Run on: Mon Mar 15 19:51:20 1999. MasPar time 8241.22 Seconds  
 Tabular output not generated. 1546 163 Million cell updates/sec

Title: >US-09-020-716-6  
 Description: (1-5392) from US09020716.seq  
 Perfect Source: 5392  
 N.A. Sequence: 1 CTAATTTGTAAGCGCTTAATA ..... ATTTCCTCGGAAGTGGCAC 5392  
 Comp: GATTTAACATTCGCAATTAT ..... TAAAGGGCGCTTTCACGGTG

Scoring table: TABLE default  
 Gap 6

Nmatch STD : Dbase 0: Query 0

Searched: 602357 seqs, 1181590623 bases x 2

Post-processing: Minimum Match 08  
 Listing first 45 summaries

Database: emb157  
 1:em\_ba 2:em\_fun 3:em\_htg 4:em\_hum1 5:em\_hum2 6:em\_in  
 7:em\_om 8:em\_or 9:em\_ov 10:em\_pat 11:em\_ph 12:em\_pl  
 13:em\_to 14:em\_un 15:em\_vi

Database: genbank110  
 16:gb\_ba1 17:gb\_ba2 18:gb\_ba3 19:gb\_in 20:gb\_om 21:gb\_ov  
 22:gb\_pat 23:gb\_ph 24:gb\_pl 25:gb\_pl2 26:gb\_pr1  
 27:gb\_pr2 28:gb\_pr3 29:gb\_ro 30:gb\_st 31:gb\_sts 32:gb\_sy  
 33:gb\_un 34:gb\_vi

Statistics: Mean 12.732 Variance 8.276, scale 1.538

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description	Pred No
1	2224	41.3	2961	32	APR2KSP	0.00e+00
2	2228	41.3	2961	32	APR2KSM	0.00e+00
3	2228	41.3	2979	32	CVKSLIC	0.00e+00
4	2228	41.3	4289	32	XXU35136	0.00e+00
5	2225	41.3	4412	32	EV05F112N	0.00e+00
6	2229	41.3	9655	32	SYNFG8V	0.00e+00
7	2229	41.3	9750	32	AF05397	0.00e+00
8	2228	41.3	9760	32	EV056815	0.00e+00
9	2229	41.3	10765	32	EV057815	0.00e+00
10	2221	41.2	2958	32	ARBLKSM	0.00e+00
11	2221	41.2	2958	32	ARBLKSP	0.00e+00
12	2224	41.2	2964	32	SYNBLKSPV	0.00e+00
13	2224	41.2	2964	32	SYNBLKSMV	0.00e+00

C	14	2224	41.2	2967	32	U02449	Cloning vector pURB01	0.00e+00
C	15	2224	41.2	7474	22	A58522	Sequence 26 from Raten	0.00e+00
C	16	2224	41.2	8072	32	AF038666	Cloning vector pFNU	0.00e+00
C	17	2223	41.0	4433	32	EV028114N	Expression vector pFNU	0.00e+00
C	18	2225	40.9	3681	22	I15651	Sequence 6 from patent	0.00e+00
C	19	2205	40.9	3681	22	A-0702	pKSEL5 DNA sequence	0.00e+00
C	20	2208	40.9	4267	32	PRS304	Yeast integrative vect	0.00e+00
C	21	2208	40.9	4373	32	PRS306	Yeast integrative vect	0.00e+00
C	22	2208	40.9	4443	32	PRS303	Yeast integrative vect	0.00e+00
C	23	2208	40.9	4783	32	PRS314	Yeast centromere vecto	0.00e+00
C	24	2228	40.9	4967	32	PRS311	Yeast centromere vecto	0.00e+00
C	25	2208	40.9	5056	32	SC064594	Yeast centromere vecto	0.00e+00
C	26	2208	40.9	5187	32	U34887	Yeast integrative vect	0.00e+00
C	27	2205	40.9	5314	22	I15642	Sequence 4 from patent	0.00e+00
C	28	2205	40.9	5314	22	A20700	FMV1 DNA sequence	0.00e+00
C	29	2208	40.9	5504	32	PRS305	Yeast integrative vect	0.00e+00
C	30	2208	40.9	5683	32	SC064593	Yeast integrative vect	0.00e+00
C	31	2208	40.9	6018	32	PRS315	Yeast centromere vecto	0.00e+00
C	32	2202	40.8	2976	32	CV061229	Cloning vector pKRX	0.00e+00
C	33	2202	40.8	3625	32	ASPGREEN1	Artificial sequences	0.00e+00
C	34	2202	40.8	3886	32	XXU35129	Plasmid pBSL141 clonin	0.00e+00
C	35	2202	40.8	4051	32	XXU35126	Plasmid pBSL121 clonin	0.00e+00
C	36	2202	40.8	4285	32	XXU35137	Plasmid pBSL199 clonin	0.00e+00
C	37	2202	40.8	4480	32	XXU35134	Plasmid pBSL190 clonin	0.00e+00
C	38	2202	40.8	4569	32	XXU35130	Plasmid pBSL142 clonin	0.00e+00
C	39	2202	40.8	4755	32	U93713	Cloning vector pRS400	0.00e+00
C	40	2202	40.8	4842	32	XXU35125	Plasmid pBSL119 eldoin	0.00e+00
C	41	2202	40.8	4893	32	U93714	Cloning vector pRS401	0.00e+00
C	42	2202	40.8	5015	32	XXU35120	Plasmid pBSL130 clonin	0.00e+00
C	43	2202	40.8	6042	32	U93718	Cloning vector pRS412	0.00e+00
C	44	2202	40.8	6042	32	U93718	Cloning vector pRS412	0.00e+00
C	45	2202	40.8	6868	32	U93719	Cloning vector pRS422	0.00e+00

ALIGNMENTS

RESULT 1  
 LOCUS APR2KSP 2961 bp DNA circular SYN 10-MAY-1995  
 DEFINITION pBluescript II KS(-) vector DNA, phagemid excised from lambda  
 ZAPII.  
 ACCESSION X59327  
 NID 958061  
 KEYWORDS artificial sequence, cloning vector, expression vector, vector,  
 cloning vectors,  
 SOURCE cloning vectors  
 ORGANISM Artificial sequences: Cloning vectors.  
 REFERENCE 1 (bases 1 to 2961)  
 AUTHORS Thomas,E.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-FEB-1990) to the EMBL/GenBank/DBJ databases, Thomas  
 E.A., Stratagene Clonin Systems, 11099 North Torrey Pines Rd., La  
 Jolla, CA 92037, USA  
 REFERENCE 2 (bases 1 to 2961)  
 AUTHORS Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.  
 TITLE Lambda ZAP: a bacteriophage lambda expression vector with in vivo  
 excision properties  
 JOURNAL Nucleic Acids Res. 16 (15), 7583-7600 (1988)  
 MEDLINE 88319944  
 REFERENCE 3 (bases 1 to 2961)  
 AUTHORS Altting-Wees,M.A. and Short,J.M.  
 TITLE pBluescript II: gene mapping vectors  
 JOURNAL Nucleic Acids Res 17 (22), 4454 (1989)  
 MEDLINE 90067967  
 FEATURES  
 Location/Qualifiers  
 1..2961  
 /organism="Cloning vector"  
 misc\_feature 1..2961  
 /note="phagemid pBluescriptII KS(-)"  
 BASE COUNT 747 a 738 c 755 g 721 t  
 ORIGIN  
 Query Match 41.3% Score 2228 DB 32 Length 2961  
 Best Local Similarity 100.0% Pred. No. 0.00e+00

[illegible][illegible]

Db 2894 CATATTTGAATGATTTAGAAAAATAAAACAAATAGGGGTTGGGCGACATTTCCCGGAAA 2953  
 QY 5325 CATATTTGAATGATTTAGAAAAATAAAACAAATAGGGGTTGGGCGACATTTCCCGGAAA 5384  
 Db 2954 AGTGGCAC 2961  
 QY 5385 AGTGGCAC 5392

RESULT 2  
 LOCUS ABL2KSM 2961 bp DNA circular SYN 10-MAY-1995  
 DEFINITION pBluescript II KS(-) vector DNA, phagemid excised from lambda ZAP11.  
 ACCESSION X52329  
 NID 958060  
 KEYWORDS artificial sequence; cloning vector; expression vector; vector.  
 SOURCE cloning vectors.  
 ORGANISM Artificial sequences; Cloning vectors.  
 REFERENCE 1 (bases 1 to 2961)  
 AUTHORS Thomas.E.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-FEB-1990) to the EMBL/GenBank/DBJ databases. Thomas E.A., Stratagene Clonin Systems, 11039 North Torrey Pines Rd., La Jolla, CA 92037, USA  
 REFERENCE 2 (bases 1 to 2961)  
 AUTHORS Short.J.M., Fernandez.J.M., Sorge.J.A. and Huse.W.D.  
 TITLE Lambda ZAP: a bacteriophage lambda expression vector with in vivo excision properties  
 JOURNAL Nucleic Acids Res. 16 (15), 7583-7600 (1988)  
 MEDLINE 88319944  
 REFERENCE 3 (bases 1 to 2961)  
 AUTHORS Alting-Mees.M.A. and Short.J.M.  
 TITLE pBluescript II: gene mapping vectors  
 JOURNAL Nucleic Acids Res. 17 (22), 9494 (1989)  
 MEDLINE 90067967  
 FEATURES  
 source Location/Qualifiers  
 1..2961  
 /organism="Cloning vector"  
 misc\_feature 1..2961  
 /note="phagemid pBluescriptII KS(-)"  
 BASE COUNT 706 a 758 c 735 g 762 t  
 ORIGIN

Query Match 41.3% Score 2228; DB 32; Length 2961;  
 Best local Similarity 100.0%; Freq No. 0.00e+00;  
 Matches 2228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 734 GTCCACCTCCAGGGGGGGGGGTTAGCCAGCTTTTGTTCCTTTTGTAGGAGTTAATTGC 793  
 QY 3165 GTCCACCTCCAGGGGGGGGGGTTAGCCAGCTTTTGTTCCTTTTGTAGGAGTTAATTGC 3224  
 Db 794 GGGCTTGGGTAATCATGGTCTATAGTGTGTTCGTTGTAATTTGTTATCCGCTCACAAAT 853  
 QY 3225 GCGTTGGGTAATCATGGTCTATAGTGTGTTCGTTGTAATTTGTTATCCGCTCACAAAT 3284  
 Db 854 TGTACACATATAGAGTGGGAGGAGTAAAGTGTAAAGGCTGGGTTGCTTAATGAGTGG 913  
 QY 3285 TCCACACATATAGAGTGGGAGGAGTAAAGTGTAAAGGCTGGGTTGCTTAATGAGTGG 3344  
 Db 914 CTAACTCACATTAATTCGGTTGGGTTGCTACCTGCGCGGTTTCCAGTGGGAAACCTGCTGG 973  
 QY 3345 CTAACTCACATTAATTCGGTTGGGTTGCTACCTGCGCGGTTTCCAGTGGGAAACCTGCTGG 3404  
 Db 974 CCAAGTGCATTAATGAATCGGCAAGCGGCGGAGAGGCGGTTTTCGTTATTTGGCGGCTC 1033  
 QY 3405 CCAAGTGCATTAATGAATCGGCAAGCGGCGGAGAGGCGGTTTTCGTTATTTGGCGGCTC 3464  
 Db 1034 TTCCGCTTCCTCGCTACTGACTGGTGGCGCTCGGCTCGGTTGCGGCTGCGGCGAGCGGTATC 1093  
 QY 3465 TTCCGCTTCCTCGCTACTGACTGGTGGCGCTCGGCTCGGTTGCGGCTGCGGCGAGCGGTATC 3524  
 Db 1094 AGCTACACAAAGGCGGTAATGAGGTTATGACAGAAATCAGGAGGATACAGGAGGAGAA 1153

QY 3525 AGCTCACATCAAGAGCGGTAAATAGGGTTATCCACAAATCAAGGTAATCAAGAAAGAA 3584  
 Db 1154 CATGTGAGCAAAAGGCCAGCAAAAGCCAGGAACCGTAAAAAGCCCGGTTGTGTGGGCTT 1213  
 QY 3585 CATGTGAGCAAAAGGCCAGCAAAAGCCAGGAACCGTAAAAAGCCCGGTTGTGTGGGCTT 3644  
 Db 1214 TTCCATAGGCTCCGCGCCCTGAGCGAGCATCAAAAAATCGAGCGCTCAAGTCAGAGGTTG 1273  
 QY 3645 TTCCATAGGCTCCGCGCCCTGAGCGAGCATCAAAAAATCGAGCGCTCAAGTCAGAGGTTG 3704  
 Db 1274 GCGAAACCCGACAGGACTATAAAGATACAGGCGGTTTCCCGCTGGAAGCTCCCTCGTGGC 1333  
 QY 3705 GCGAAACCCGACAGGACTATAAAGATACAGGCGGTTTCCCGCTGGAAGCTCCCTCGTGGC 3764  
 Db 1334 CTCTCGTGTTCGAGCGCTTACCGGCTTACCGGATACAGTGTCCGCTTCTCCCTTCGGGAA 1393  
 QY 3765 CTCTCGTGTTCGAGCGCTTACCGGCTTACCGGATACAGTGTCCGCTTCTCCCTTCGGGAA 3824  
 Db 1394 CGTGGCGCTTCTCATAGCTACGCTGTAGGTATCTCAAGTTCGGGTGAGTGTCTGCTG 1453  
 QY 3825 CGTGGCGCTTCTCATAGCTACGCTGTAGGTATCTCAAGTTCGGGTGAGTGTCTGCTG 3894  
 Db 1454 CAAGTGTGGCTGTGTGCAAGCAACCCCGGTTTACGCTGCAAGCGGCTTTCGCTTATCGGTA 1513  
 QY 3885 CAAGTGTGGCTGTGTGCAAGCAACCCCGGTTTACGCTGCAAGCGGCTTTCGCTTATCGGTA 3944  
 Db 1514 CTATCGTGTGAGTCCCAACCCCGGTTAAGACAGCACTTATCGGCTAGTGGAGTACGCA 1573  
 QY 3945 CTATCGTGTGAGTCCCAACCCCGGTTAAGACAGCACTTATCGGCTAGTGGAGTACGCA 4004  
 Db 1574 TAACAGGATTAAGCAGAGTATGATAGGCTGTATAGAGTTCCTTGAAGTGTGGTGGC 1633  
 QY 4005 TAAACAGGATTAAGCAGAGTATGATAGGCTGTATAGAGTTCCTTGAAGTGTGGTGGC 4064  
 Db 1634 TAACTAGGCTTACACTAGAGGACAGATTTGGTATCTGGGTTCTGTGTGAAGCAAGTAC 1693  
 QY 4065 TAACTAGGCTTACACTAGAGGACAGATTTGGTATCTGGGTTCTGTGTGAAGCAAGTAC 4154  
 Db 1694 CTTCGGAAGAGGTTGGTGTAGTCTTGTATCGGCAACCAACCAAGCGTGTAGGAGTGG 1753  
 QY 4125 CTTCGGAAGAGGTTGGTGTAGTCTTGTATCGGCAACCAACCAAGCGTGTAGGAGTGG 4184  
 Db 1754 TTTTGTGTTCGAAGCAGCAGATTAAGCGCAGAAAAAAATGATGCAAGAAAGATCTTT 1813  
 QY 4185 TTTTGTGTTCGAAGCAGCAGATTAAGCGCAGAAAAAAATGATGCAAGAAAGATCTTT 4244  
 Db 1814 GATCTTTTCTAGGGGTCTGAGCTCAGTGGCAACCAAACTCAGTTAAGGAGTTTGGT 1873  
 QY 4245 GATCTTTTCTAGGGGTCTGAGCTCAGTGGCAACCAAACTCAGTTAAGGAGTTTGGT 4304  
 Db 1874 CATGAGATTATCAAAAGGATCTTCACCTAGATCTTTTAAATTAATAAATGAAGTTTAA 1933  
 QY 4305 CATGAGATTATCAAAAGGATCTTCACCTAGATCTTTTAAATTAATAAATGAAGTTTAA 4364  
 Db 1934 ATCAATCTAAAGTATATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAAATCAGTGA 1993  
 QY 4365 ATCAATCTAAAGTATATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAAATCAGTGA 4424  
 Db 1994 GGCACCTATCTCAGCGATCTGTCTATTTTGGTTCATCCATAGTTGGCTGCTCCCGGCTG 2053  
 QY 4425 GGCACCTATCTCAGCGATCTGTCTATTTGGTTCATCCATAGTTGGCTGCTCCCGGCTG 4484  
 Db 2054 GTAGATAACTACGATACGGGAGGGGTTACCATCTGTGAGGAGGAGTGTGAAATGATAGCG 2113  
 QY 4485 GTAGATAACTACGATACGGGAGGGGTTACCATCTGTGAGGAGGAGTGTGAAATGATAGCG 4544  
 Db 2114 AGACCCACGCTCAGCGGCTCCAGATTTATACAAATAAACAACCAAGGAGGAGGAGTGA 2173  
 QY 4545 AGACCCACGCTCAGCGGCTCCAGATTTATACAAATAAACAACCAAGGAGGAGGAGTGA 4604  
 Db 2174 GCGCAGAGTGTCTCTGCAACTTTATCCGCTTCCATCCAGTCTATTAATTTGTTCCGAGGA 2233







```

4485 GTAGATAAATACGATACGGAGGGGCTTACATATGAGGACAGTGTCTCAATATATACGGG 4544
Db 3442 AGACGTAACGCTACCGGCTCCAGATTATACAGCAATAAACCCAGCAGCCGGAAGGGCGGA 3501
QY 4545 AGACGTAACGCTACCGGCTCCAGATTATACAGCAATAAACCCAGCAGCCGGAAGGGCGGA 4604
Db 4502 AGACGTAACGCTACCGGCTCCAGATTATACAGCAATAAACCCAGCAGCCGGAAGGGCGGA 3561
QY 4605 AGACGTAACGCTACCGGCTCCAGATTATACAGCAATAAACCCAGCAGCCGGAAGGGCGGA 4664
Db 3562 AGCTAGTAAGTAGTTCGCCAGTTAATAGTTTGGCGAACGTTGTTGCCATTGTCACAGG 3621
QY 4665 AGCTAGTAAGTAGTTCGCCAGTTAATAGTTTGGCGAACGTTGTTGCCATTGTCACAGG 4724
Db 3622 CATCGTGTGTGACGCTGCTGCTGTTGGTATGCTTCATTCAGCTCCGCTTCCACAGCATC 3681
QY 4725 CATCGTGTGTGACGCTGCTGCTGTTGGTATGCTTCATTCAGCTCCGCTTCCACAGCATC 4784
Db 3682 AAGGCGAGGTACATATGACGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3741
QY 4785 AAGGCGAGGTACATATGACGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4844
Db 3742 GATGCTGTGTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3801
QY 4845 GATGCTGTGTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4904
Db 3802 TAATCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3861
QY 4905 TAATCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4964
Db 3862 TAATCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3921
QY 4965 TAATCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5024
Db 3922 GATGCTGTGTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3981
QY 5025 GATGCTGTGTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5084
Db 3982 GGGGCGGAAACTCTCAAGGATGCTTACCGCTGTTTGAGATCCAGTTCGATGTAAACCCACTCG 4041
QY 5085 GGGGCGGAAACTCTCAAGGATGCTTACCGCTGTTTGAGATCCAGTTCGATGTAAACCCACTCG 5144
Db 4042 TGACGCAACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4101
QY 5145 TGACGCAACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5204
Db 4102 AGAAGGCGAAATGCGGCGGAAAGGGAATAAGGGCGACAGGGAATGTTGAATACATCAT 4161
QY 5205 AGAAGGCGAAATGCGGCGGAAAGGGAATAAGGGCGACAGGGAATGTTGAATACATCAT 5264
Db 4162 ACTCTTCCTTTTCAATATTATGAGCAATTTATCAGGCTTATGCTGCTCATGAGCGGATA 4221
QY 5265 ACTCTTCCTTTTCAATATTATGAGCAATTTATCAGGCTTATGCTGCTCATGAGCGGATA 5324
Db 4222 CATATTTGAATGTTATGAGCAATTTATCAGGCTTATGCTGCTCATGAGCGGATA 4281
QY 5325 CATATTTGAATGTTATGAGCAATTTATCAGGCTTATGCTGCTCATGAGCGGATA 5384
Db 4282 AGTGGCAC 4289
QY 5385 AGTGGCAC 5392

RESULT 5
LOCUS EVCOR112N 4412 bp DNA SYN 18-MAR-1996
DEFINITION Expression vector pCOR112N (modified from pCOR112 in [2]).
ACCESSION 229587
ID 4452344
KEYWORDS ampicillin resistance; beta-lactamase; ColE1 origin; expression
vector; nos terminator; phage fl region; rice actin1 promoter.
SOURCE expression vectors
ORGANISM expression vectors

```

```

artificial sequence.
1 (bases 1 to 4412)
Liu, L., Basqupta, I., Davies, J.W. and Hull, R.
Modified vectors for monocot transformation toward virus resistance
Unpublished
REFERENCE 2 (bases 1 to 4412)
McElroy, D., Blowers, A.D., Jones, B. and Wu, R.
Construction of expression vectors based on the rice actin 1 (Act1)
5' region for use in monocot transformation
Mol. Gen. Genet. 231 (1), 150-160 (1991)
JOURNAL 92092956
MEDLINE 92092956
REMARK 3 (bases 27 to 282)
Depicker, A., Stachel, S., Phares, P., Zaitsev, P. and Goodman, H.M.
Nopaline synthase: transcript mapping and DNA sequence
J. Mol. Appl. Genet. 1 (6), 561-573 (1982)
JOURNAL 83110651
MEDLINE 83110651
REFERENCE 4 (bases 1 to 4412)
Liu, L.
Direct Submission
Submitted (26-JAN-1994) Liu L., John Innes Institute, Virus
Research, Colney Lane, Norwich, United Kingdom, NR4 7JH
JOURNAL
FEATURES
Source
Location/Qualifiers
1..4412
/organism="expression vectors"
/db_xref="taxon:33776"
misc_feature 1..30
/note="multiple cloning sites"
misc_feature 3..5
/note="start codon for gene expression"
terminator 27..282
/note="nos terminator"
/citation={3}
506..962
/note="phage fl region"
1093..1953
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/db_xref="PID:g452345"
/transl_table="MSIHPFVALIDFEFAAFPLVFAHPBELVKKAEFOLGAEVSV
TELDNSKILSEFPPEERPMSTFKVITGSAVSPFAGTQELPPFIRISQNCLE
YSPVTEKHLIDGMTVRECCSAATMSDNTAANLLTTIGPKELIAFPHMDHVTPL
DPPELENAIPNDPOTTPVAMATLPLTITLILASPOOLDMFAOKVAGPL
LRSALPAGWFIADKSGAGRGSPGIIAALGPKGKPSPIVYITGQATMDENRQIA
EIGASLIKHW"
misc_feature 1954..2894
/note="ColE1 origin"
3181..4412
/note="rice actin 1 5' region (act1 promoter)"
/citation={2}
BASE COUNT 1127 a 1107 c 1036 g 1142 t
ORIGIN
Query Match 41.38% Score 2225; DB 32; Length 4412;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 965 GTGGCACATTTTCGGGGAATGTGGCGGAACCCCTATTGTTTATTTCTTAATAATATT 1024
Cp 5392 GTGGCACATTTTCGGGGAATGTGGCGGAACCCCTATTGTTTATTTCTTAATAATATT 5333
Db 1025 CAATATGATGCTGCTCATGAGCAATAACCTGATAAATGCTTCAATATATGAAAAA 1084
Cp 5332 CAATATGATGCTGCTCATGAGCAATAACCTGATAAATGCTTCAATATATGAAAAA 5273
Db 1085 GGAAGAGTAGTAGTATTCACCAATTTCCGTGCGGCCCTATTCCCTTTTTCGGCATTTT 1144
Cp 5272 GGAAGAGTAGTAGTATTCACCAATTTCCGTGCGGCCCTATTCCCTTTTTCGGCATTTT 5213
Db 1145 GCTTCCTGTTTTTCTCTACCCAGAAAGCTTGGTGAAGATGCTGAAGATAGT 1204
Cp 5212 GCTTCCTGTTTTTCTCTACCCAGAAAGCTTGGTGAAGATGCTGAAGATAGT 5153

```







```

artificial sequence; cloning vectors.
1 (bases 1 to 9655)
AUTHORS Mazzarella, R. and Pillia, G.
TITLE Recombination trapping: an 'in vivo' approach to recover cDNAs
        encoded in YACS
JOURNAL Unpublished (1997)
COMMENT GSDB:S:1274444.
FEATURES
    Location/Qualifiers
        source
            1..9655
                /organism="Cloning vector pGR8"
                /plasmid="Cloning vector pGR8"
                /db_xref="taxon:64712"
                1..135
                    /standard_name="telomere"
        misc_feature
            145..2097
                /gene="HIS5"
            145..2097
                /gene="HIS5"
            761..1915
                /gene="HIS5"
            /EC_number="2.6.1.9"
            /codon_start=1
            /transl_table=11
            /product="histidinal phosphate aminotransferase"
            /db_xref="PID:g2338009"
            /translation="MVFDLRIVRPKIYNLEPYRCARDFTFEGILLDANENAHGPTPV
            ELKTLNHRYPDPHOLEFKIAMTKYFNKTSYVANDPEVKPLTADNLCGLGVSDSIDA
            IIPACCVREKILVLPPYSSMSSVCANINDIEVCCPLIVSDGSFQMDTEAVLTILKN
            DSIKLMFVPGNTGAKIKTSLIEKVLQWNGLVVDAYVDYDFGGSTAPLVTKY
            PNLVQTLKSGFAGIKLGMIAELARILNMAKAPINISSLALEYALKAVQDSN
            LKMEATSIINEEKLLKLTALDYDDQVGGDLNELLIRINGDQNVLAKKLYV
            QLATQGVVVRFRGNELCGCLRTITVGTHEENTHLIKYFKETLYKLANE"
        prim_transcript
            2138..6708
                /gene="LYS2"
            2138..6708
                /gene="LYS2"
            2374..6552
                /gene="LYS2"
            /EC_number="1.2.1.31"
            /codon_start=1
            /transl_table=11
            /product="alpha-aminoadipate reductase"
            /db_xref="PID:g2358310"
            /translation="MTNEKWIIEKLDNPTLSLPHDFLRPOQEPYTKQATYSVLQPL
            DVPHDSFNKVAVALSVMAALIYRVGTGDDIVLYIANKNLIKRFNIQPTWSELYSTI
            NENLKNLSIEANFSDELAEKIQSCDLERTPOLFRALFLENODFKLDEKHLHVD
            ALNDTSNAHVLNLIYNSLYSNERVTLVADQFTQYLTAALSDPNCIKLISLTAS
            SKDSLDPKTLNKGWDFVCCIHDIEQDNAEAPPERTCVETPTLNSKRSFTTRDIN
            FTSNVAHTLCIKTIGKIPGSDVMYIYSSPGVDLMVCVNGVKAGATFSVIDPAPPAROT
            IYLGVAKPLGLIVRAAGQDLQVEDYINDELEIVSRINSIAIQENGTIEGKLNDGE
            DVLAPEYDKDTIRTVVVGPDNPTLSFTSGSEIPKGVLRHFSLAYEWNMSKREN
            LTENDKFTLSGIAHDPIDQDMFTPLFLGAOLYVPTQDDIGTPGRLAEMWSKYCTVT
            HLTTPMGQLTLQAATPPPKLHAFVGDILTFRDLRLQTLAENCRIVMNGYGTETQ
            RAVSEYFKNVDNDNFKLLKDVMPACKGMLNVLQVLLVNRDRTQICGIEIGIYV
            RAGLAEGYKGLPELNKEFVNNEVEKDHNYLDNKGPEWRFQWGLRDNRTDGT
            LGRYPNGCECCGADDDQVKIRGFEIEGEIDTHISQHLVRENITLVRKNADNFT
            LITFMVYMDKPLDKSQFSDQFKEVEDPIVKGLIYHLLSKDIRFLPKRLASAYAM
            PSILVYMDKPLNPNKVDKPKLQPTPKQLNVAENTVSETDSDQFTNVEREVDLW
            LSLPTKPAVSPPDSQFVFDLGSILATKMLFTLKKLQVLDPLGTFPKYPTIKAFPA
            EIDRIKSSGSSQGEVENVNTANYAEADKLVETLPSSYPSPREFVFPNSAGKTTIN
            VFVTGVTGLSYILADLLGRSPKNYSFKFAHVRADKEEAFARQKLTGTWNE
            KFAINSVLVGLDQSKSQFGLSDEKMDLANTVDIIHNGALVHWYYPYAKLRDPNVIS
            TINWMLAAGKPKDFEVSSTILDTEYFNLSKLVSEKPGIILESDLLMSASGL
            TGGYGSQKWAAYIIRRAGERGLRGCIVRPGYVTVGASANGSNTDDFLRLPKLSVOL
            GKIPDTENSVMVVPDVARVVVATSLNPPKENELAVQVTHGPHRIILFKDYLYLTHDY
            GYDVEIESYKMKKSLSEASVIDRNEENALYPLLHMVDNLPESTKAPELDRNVAVL
            KKDANTGVDSWNGICVTPPEEVGIVIAFLNKVGLFPPTTHNDKPLPSIELTQAISL
            VASGAGARSSAAA"
        misc_feature
            complement(6709..7185)
        gene
            7186..7496
                /gene="Lac2"
        prim_transcript
            7186..7496

```

```

/misc_feature
    7339..7445
        /gene="Lac2"
    /standard_name="multiple cloning sites"
    /complement(7910..8030)
    /standard_name="plasmid ORI"
    /complement(8659..9522)
    /gene="AMP"
    /complement(8659..9519)
    /gene="AMP"
    /function="ampicillin resistance"
    /codon_start=1
    /transl_table=11
    /product="beta-lactamase"
    /db_xref="PID:g2358011"
    /translation="MSIQHFRVALIPFAAFCLPVPFAHPETLVKVKDAEDQLGARVGY
    IELDNSKILLESFRRPEFMFTFKVILGGAVLSLTDAQGGGLLQAPRRHYSNLSLVE
    YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLITIGGKELTFLHNMGDHVTFL
    DRWPELNEAIPNDERUTIMPVAMATILRLKLLTGELLILASRQQLIDWMEADKVAQPL
    LPSALPAGWFIADKSGAGEPSGIIAALGPDGKPSPIVVIYITGSGATWCEENPQIA
    EIGASLIKHW"
    prim_transcript complement(8734..9522)
    /gene="AMP"
BASE COUNT      2804 a 2009 c 2099 g 2743 t
ORIGIN
Query Match      41.3%   Score 2228;   DB 32;   Length 9655;
Best Local Similarity 100.0%;   Pred. NO. 0.00e+00;
Matches 2228;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;
Db 7420  GTGAGACTCGAGGGGGGGGATCCAGAGTITTTTTCCTTTTATAGTGGAGGTTAAATTCG 7479
Qy 3165  GTGAGACTCGAGGGGGGGGATCCAGAGTITTTTTCCTTTTATAGTGGAGGTTAAATTCG 3224
Db 7480  GCCTTGGGGTAAATCAATGATAGTCTGTTTCTCTGTGTAATTTCTTATCGCTCAATAAT 7539
Qy 3225  GCCTTGGGGTAAATCAATGATAGTCTGTTTCTCTGTGTAATTTCTTATCGCTCAATAAT 3284
Db 7540  TCCACACACATACAGAGCGGAGGATAAAGTGTAAAGCGTGGGGTGGGTCGACCAAGAGTAG 7599
Qy 3285  TCCACACACATACAGAGCGGAGGATAAAGTGTAAAGCGTGGGGTGGGTCGACCAAGAGTAG 3344
Db 7600  CTAATCTACATTAATTTGCTTTCCTTCATCTGCTGCTGCTTTCACAGTGGGAAACGCTGGTGG 7659
Qy 3345  CTAATCTACATTAATTTGCTTTCCTTCATCTGCTGCTGCTTTCACAGTGGGAAACGCTGGTGG 3404
Db 7660  CCAGCTGCATTAAATGAATCGGCAACGGCGGGGAGAGGGCGTTTTCGCTATTTGGCGCTC 7719
Qy 3405  CCAGCTGCATTAAATGAATCGGCAACGGCGGGGAGAGGGCGTTTTCGCTATTTGGCGCTC 3464
Db 7720  TTCGCTTTCCTGCTCACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7779
Qy 3465  TTCGCTTTCCTGCTCACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3524
Db 7780  AGTCTACTCAAGGGCGGTAAATCGGTTATTCACAAATCAGGSGGATAACAGCAGGAAAGAA 7839
Qy 3525  AGTCTACTCAAGGGCGGTAAATCGGTTATTCACAAATCAGGSGGATAACAGCAGGAAAGAA 3584
Db 7840  CATGTGAGCAAAAGGCCAGCAAAAGGCCAGCAAAAGGCCAGCAAAAGGCCAGCAAAAGGCCAG 7899
Qy 3585  CATGTGAGCAAAAGGCCAGCAAAAGGCCAGCAAAAGGCCAGCAAAAGGCCAGCAAAAGGCCAG 3644
Db 7900  TTTCCTATAGGCTCCGCGCGCTGCTGAGTACATCAAAAATGACGCTCAAGTTCAGAGGCTG 7959
Qy 3645  TTTCCTATAGGCTCCGCGCGCTGCTGAGTACATCAAAAATGACGCTCAAGTTCAGAGGCTG 3704
Db 7960  GCGAAACCCGACAGGACTATAAGATACACAGCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCG 8019
Qy 3705  GCGAAACCCGACAGGACTATAAGATACACAGCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCG 3764
Db 8020  CTCTCTCTGTTCCGACCGCTGCGCTTACCGGATACCTGTGCGGCTTTCCTTTCCTTTCCTTTC 8079
Qy 3765  CTCTCTCTGTTCCGACCGCTGCGCTTACCGGATACCTGTGCGGCTTTCCTTTCCTTTCCTTTC 3824

```



Dh 1393 GGAAGATATGATTAACATTTCCGTGTCGCCCTTATTCCTTTTTCGGCATTTT 1452  
Cp 5272 GGAAGATATGATTAACATTTCCGTGTCGCCCTTATTCCTTTTTCGGCATTTT 5213  
Dh 1453 GCCTTCCTGTTTTCCTACCCAGAACCGTGTGAAAGTAAAGATGCTGAAGATCAGT 1512  
Cp 5212 GCCTTCCTGTTTTCCTACCCAGAACCGTGTGAAAGTAAAGATGCTGAAGATCAGT 5153  
Dh 1513 TGGGTGACAGTGGTGTATCATCGAATCGAATCGAATCGAATCGAATCGAATCGAAT 1572  
Cp 5152 TGGGTGACAGTGGTGTATCATCGAATCGAATCGAATCGAATCGAATCGAATCGAAT 5093  
Dh 1573 TTCCGCCGGAAGAACGTTTTCCTAATGATGAGCAGTATTTAAAGTTTCCTGATGCGCGG 1632  
Cp 5092 TTCCGCCGGAAGAACGTTTTCCTAATGATGAGCAGTATTTAAAGTTTCCTGATGCGCGG 5033  
Dh 1633 TATTATCCCTATGAGCGCGGCAAGAGCACTCGGTGCGCGCATACACTATTCCTAGA 1692  
Cp 5032 TATTATCCCTATGAGCGCGGCAAGAGCACTCGGTGCGCGCATACACTATTCCTAGA 4973  
Dh 1693 ATGACTTGGTGTAGTATCACTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 1752  
Cp 4972 ATGACTTGGTGTAGTATCACTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 4913  
Dh 1753 GAGAAATATGAGTGTGATTAACCATGATGATGATGATGATGATGATGATGATGAT 1812  
Cp 4912 GAGAAATATGAGTGTGATTAACCATGATGATGATGATGATGATGATGATGATGAT 4853  
Dh 1813 CAACGATGAGGAGCGAAGAGCACTAACGCTTTTTCGACACATGGGGATCATGTAA 1872  
Cp 4852 CAACGATGAGGAGCGAAGAGCACTAACGCTTTTTCGACACATGGGGATCATGTAA 4793  
Dh 1873 CTGCGCTTGTATGTTGGAACGAGAGTGAATGAAGCCATACCAACGACGAGCGTGACA 1932  
Cp 4792 CTGCGCTTGTATGTTGGAACGAGAGTGAATGAAGCCATACCAACGACGAGCGTGACA 4733  
Dh 1933 CCAGATGAGTGTAGTAAATGGCAACAGCTTGGCGAAACTATTAAGTGGCAACTACTTA 1992  
Cp 4732 CCAGATGAGTGTAGTAAATGGCAACAGCTTGGCGAAACTATTAAGTGGCAACTACTTA 4673  
Dh 1993 CTCTAGCTTCCGCGCAACATTAATAGACTGGATGAGGCGGATGAAGTTGCGAGCAC 2052  
Cp 4672 CTCTAGCTTCCGCGCAACATTAATAGACTGGATGAGGCGGATGAAGTTGCGAGCAC 4613  
Dh 2053 TTCTGCGCTGCGGCTTTCGCGCTGCGTGTATTTATGCTATTAATCTGAGCGGTGAGC 2112  
Cp 4612 TTCTGCGCTGCGGCTTTCGCGCTGCGTGTATTTATGCTATTAATCTGAGCGGTGAGC 4553  
Dh 2113 GTGGGTCTGCGGTATCATTTGAGCACTGGGCGGATGATGGTAAAGCTTCCGTATCGTAG 2172  
Cp 4552 GTGGGTCTGCGGTATCATTTGAGCACTGGGCGGATGATGGTAAAGCTTCCGTATCGTAG 4493  
Dh 2173 TTATCTACAGCAGGAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 2232  
Cp 4492 TTATCTACAGCAGGAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 4433  
Dh 2233 TAGTGCTGCTACTGATTAAGCAATGTAAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGT 2292  
Cp 4432 TAGTGCTGCTACTGATTAAGCAATGTAAGTGTGAGCAAGTGTGAGCAAGTGTGAGCAAGT 4373  
Dh 2293 AGATTGATTTAAACTTCACTTTTAAATTTAAAGGATCTAGGTGAAGATCCCTTTTGATA 2352  
Cp 4372 AGATTGATTTAAACTTCACTTTTAAATTTAAAGGATCTAGGTGAAGATCCCTTTTGATA 4313  
Dh 2353 ATCTCATGACCAAAATCCCTTAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2412  
Cp 4312 ATCTCATGACCAAAATCCCTTAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 4253  
Dh 2413 AAAAGATCAAGGATCTTCTTGAGATCCCTTTTTCGCGGTAATCTGCTGCTTGCAAA 2472  
Cp 4252 AAAAGATCAAGGATCTTCTTGAGATCCCTTTTTCGCGGTAATCTGCTGCTTGCAAA 4193  
Dh 2473 CAAAAAACCACCTTACAGCGTGGTGTGTTGTCGCGATCAAGAGTACCAACTCTTT 2532

Cp 4192 CAAAAAACCACCTTACAGCGTGGTGTGTTTGTGCGGATCAAGAGTACCAACTCTTT 4133  
Dh 2533 TTCCGAAGTACTGCTTTCAGCAGAGCGCAGATACCAAACTACTGCTCTTCTAGTGTAGC 2592  
Cp 4132 TTCCGAAGTACTGCTTTCAGCAGAGCGCAGATACCAAACTACTGCTCTTCTAGTGTAGC 4073  
Dh 2592 CGTAGTATGCGCACACATTCAGAACTCTGTAAGCAGCGGATACATACCTGCTCTGTAA 2652  
Cp 4072 CGTAGTATGCGCACACATTCAGAACTCTGTAAGCAGCGGATACATACCTGCTCTGTAA 4013  
Dh 2652 TCCTGTTACAGTGGCTGCTGCACTGCTGTAAGTGTGCTTACCGGTTGACACAA 2712  
Cp 4012 TCCTGTTACAGTGGCTGCTGCACTGCTGTAAGTGTGCTTACCGGTTGACACAA 3953  
Dh 2712 GAGCATAGTACCGGATAGCGGCGAGCGTGGGCTGACCGGGGCTTCCTGCAACAGC 2772  
Cp 3952 GAGCATAGTACCGGATAGCGGCGAGCGTGGGCTGACCGGGGCTTCCTGCAACAGC 3893  
Dh 2772 CCAGCTTGGAGCGAAGCACTACAGCACTGAGATACCTACAGCGTGAATGASAA 2832  
Cp 3892 CCAGCTTGGAGCGAAGCACTACAGCACTGAGATACCTACAGCGTGAATGASAA 3833  
Dh 2833 GCGCCAGCTTCCGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2892  
Cp 3832 GCGCCAGCTTCCGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3773  
Dh 2892 CAGCAGTGCAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2952  
Cp 3772 CAGCAGTGCAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3713  
Dh 2952 GGTTCGCCACCTCTGACTTGGCGTTCGATTTTGTGATGCTGCTGAGGCGGCGAGGCG 3012  
Cp 3712 GGTTCGCCACCTCTGACTTGGCGTTCGATTTTGTGATGCTGCTGAGGCGGCGAGGCG 3653  
Dh 3012 TATGGAAGAACGCGCAGCGGCGCTTTTACGGTTCCTGCGCTTTTCTGCGCTTTT 3072  
Cp 3652 TATGGAAGAACGCGCAGCGGCGCTTTTACGGTTCCTGCGCTTTTCTGCGCTTTT 3593  
Dh 3072 CTACATGTTCTTCTGCGTTCCTGCTGATTTTGTGATGCTGCTGAGGCGGCGAGGCG 3132  
Cp 3592 CTACATGTTCTTCTGCGTTCCTGCTGATTTTGTGATGCTGCTGAGGCGGCGAGGCG 3533  
Dh 3132 AGTGAGTGTATACCGTTCGCGCAGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 3192  
Cp 3532 AGTGAGTGTATACCGTTCGCGCAGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 3473  
Dh 3192 AAGCGGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3252  
Cp 3472 AAGCGGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3413  
Dh 3252 GCAGCTGGCAGCAGCGGCTTTCGCGCTGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGG 3312  
Cp 3412 GCAGCTGGCAGCAGCGGCTTTCGCGCTGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGG 3353  
Dh 3312 TGAGTGTACTCTATAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3372  
Cp 3352 TGAGTGTACTCTATAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3293  
Dh 3372 TGCTGGAATTTGAGCGGATACAAATTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3432  
Cp 3292 TGCTGGAATTTGAGCGGATACAAATTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3233  
Dh 3432 CCAAGCGCGCAATTAACCTTCACTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3492  
Cp 3232 CCAAGCGCGCAATTAACCTTCACTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3173  
Dh 3492 AGGTGAGC 3500  
Cp 3172 AGGTGAGC 3165

RESULT 8

LOCUS EVR66815 9760 bp DNA SYN 17-DEC-1997  
 DEFINITION Expression vector pESP-2 beta lactamase and beta isopropylmalate  
 dehydrogenase genes, complete cds.  
 ACCESSION U86815  
 NID 42692339  
 KEYW PDS  
 SOURCE  
 ORGANISM  
 Expression vector pESP-2.  
 Expression vector pESP-2.  
 artificial sequence, expression vectors.  
 REFERENCE  
 1 (bases 1 to 9760)  
 Lu.Q., Bauer, J.C., and Greener, A.  
 Using Schizosaccharomyces pombe as a host for expression and  
 purification of eukaryotic proteins  
 Gene 200 (1-2), 135-144 (1997)  
 JOURNAL  
 MEDLINE  
 98038984  
 REFERENCE  
 2 (bases 1 to 9760)  
 Lu.Q.  
 Direct Submission  
 Submitted (23-JAN-1997) RefSeq, StrataGene, 11011 North Torrey Pines  
 Road, La Jolla, CA 92037, USA  
 JOURNAL  
 3 (bases 1 to 9760)  
 Lu.Q.  
 Direct Submission  
 Submitted (17-DEC-1997) RefSeq, StrataGene, 11011 North Torrey Pines  
 Road, La Jolla, CA 92037, USA  
 JOURNAL  
 REMARK  
 Sequence update by submitter  
 On Dec 17, 1997 this sequence version replaced ai:1842041.  
 FEATURES  
 Location/Qualifiers  
 1..9760  
 /organism="Expression vector pESP-2"  
 /specific\_host="Schizosaccharomyces pombe"  
 /db\_xref="taxon:56268"  
 1401..2261  
 /codon\_start=1  
 /product="beta-lactamase"  
 /db\_xref="F016:J1842042"  
 2262..340  
 /translation="MSIQHFAVLIPFAAFCLPVFAHPETLVKVKDAEDQLGARVGY  
 FELDKSKIIESPPEEPFMSTFEVLIIGAVISPTAGSFCQGFTHYSCNDLVE  
 YSPVTEKHLIDGMIVRELCSAAITMSDNTAANLLTIIGPKELTAFLEHMDGVTRL  
 DRWPELNEAIPNDRTTPVAMATILRLKLLIGELLILASRQQLIDNMEADKVGPL  
 LRSALPAGMFIADKSGAGRGSRCLIAALCPNGKPSPIVVIYTTSSQATMDPEPQIA  
 ELGASLIKHW"  
 complement(3912..5006)  
 /codon\_start=1  
 /product="beta-isopropylmalate dehydrogenase"  
 /db\_xref="F016:J2692340"  
 /translation="MSAPKIVLVLPDHWGQBITAEAIKVLKALSDVRSNVKDFENH  
 LIGCAATATGVLIPDEAIKSKKVAVLIGANGPKWGTQVVPPEQGLLKIPKELQL  
 YANLRPNFASDSILGLSPIKPKLFAKGIIEVVVRELVGSIYFKRRKEDDGDGVMSD  
 GYTVPEVQITRMAAFMALQHEPPLPIWLSKANKVILASSRLKRTVEETIKRPFPTLK  
 VQHLIDSAAMILVKNPTHLNGLIITSNMFGDIISDEASVIGSLGLPSLASLPLD  
 KNTAFGLYEPCHGSAPLPKPKNKPATILISAAMMLKLSINLPEEKATFPAVRKRVLD  
 AGIRTGDLGSGNSNTEYGDVAEEVKLLA"  
 BASE COUNT 2020 a 1962 c 1983 g 2805 t  
 ORIGIN  
 Query Match 41.3% Score 2228; DB 32; Length 9760;  
 Best Local Similarity 100.0%; Pred. No. G:30, A:02;  
 Matches 2229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1273 GTGCTATTTTGGGAAATGGGAGAAAGCTATTTGTTATTTTTCGAAATACATT 1332  
 Cp 6392 GTGCTATTTTGGGAAATGGGAGAAAGCTATTTGTTATTTTTCGAAATACATT 5333  
 Db 1433 CAATATGTATCGGTATGACACATAAAGCTGATAAAGCTTCATATATTTGAAAAA 1392  
 Cp 6332 CAATATGTATCGGTATGACACATAAAGCTGATAAAGCTTCATATATTTGAAAA 5273  
 Db 1493 GCAAGACTATGATTTAAATTTTGGTGTGCTGCTATTCCTTTTTCGGGCAATTT 1452  
 Cp 6272 GCAAGACTATGATTTAAATTTTGGTGTGCTGCTATTCCTTTTTCGGGCAATTT 5213  
 Db 1453 GCGTTCTCTTTTGTCTACACAAAGAGCTGTGTCAAAATAAAGAGATGCTGAAATCAGT 1512

Cp 5212 GCTTCCTGTTTTTGGTCAAGGAGAAAGCTGTGAAAGTAAAGAGTGAACATPACT 5193  
 Db 1513 TGGGTACAGTASTGTTTATATGAACTGGATCTCAACAGTGGTAAATCTTTGAAAT 1472  
 Cp 5152 TGGGTGACAGTASTGTTTATATGAACTGGATCTCAACAGTGGTAAATCTTTGAAAT 5093  
 Db 1573 TGGGTGAGAGAAAGTGGTGGTAAAGAGTAAATCTTTGAAATCTTTGAAAT 1532  
 Cp 5092 TGGGTGAGAGAAAGTGGTGGTAAAGAGTAAATCTTTGAAATCTTTGAAAT 5053  
 Db 1633 TATATATGCTATATGAGTGGGAAAGAGTAAATCTTTGAAATCTTTGAAAT 1592  
 Cp 5032 TATATATGCTATATGAGTGGGAAAGAGTAAATCTTTGAAATCTTTGAAAT 4973  
 Db 1693 AIGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1652  
 Cp 4972 AIGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 4933  
 Db 1753 GAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1712  
 Cp 4912 GAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 4873  
 Db 1813 CAATGATGAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAG 1772  
 Cp 4852 CAATGATGAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAG 4793  
 Db 1873 CTGGTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1832  
 Cp 4792 CTGGTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 4753  
 Db 1933 CTAGATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1892  
 Cp 4732 CTAGATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 4693  
 Db 1993 CTGTATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1952  
 Cp 4672 CTGTATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 4633  
 Db 2053 TTCTGGCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGG 2012  
 Cp 4612 TTCTGGCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGG 4573  
 Db 2113 GTGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGG 2072  
 Cp 4552 GTGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGG 4513  
 Db 2173 TTATCTACAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAG 2132  
 Cp 4492 TTATCTACAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAG 4453  
 Db 2233 TAGTGTCTCTACTGATTAAGCTGATTAAGCTGATTAAGCTGATTAAGCTGATTA 2192  
 Cp 4432 TAGTGTCTCTACTGATTAAGCTGATTAAGCTGATTAAGCTGATTAAGCTGATTA 4393  
 Db 2293 AGATTGATTTAAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2252  
 Cp 4372 AGATTGATTTAAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4333  
 Db 2353 ATCTCATGACCAAAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2312  
 Cp 4312 ATCTCATGACCAAAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4273  
 Db 2413 AAAAGATCAAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2372  
 Cp 4252 AAAAGATCAAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4213  
 Db 2473 CAAAAAACCACCGCTACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2432  
 Cp 4192 CAAAAAACCACCGCTACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4153  
 Db 2533 TTCTGCAAGTAAATGGGCTGCAAGTAAATGGGCTGCAAGTAAATGGGCTGCAAGT 2492

RESULT	9	
LOCUS	EVU67875	10765 bp
DEFINITION	PESP-1 yeast expression vector, complete sequence.	DNA circular SYN
ACCESSION	U57875	
KEYWORDS	g2589262	
	NID	
	KEYWORDS	

Db	1273	GTGGGACATTTTCGGGGAATGTGGCGGAACCCCTATTGTTTATTGTTTCTTAATACATT	1332
Cp	5392	GTGGGACATTTTCGGGGAATGTGGCGGAACCCCTATTGTTTATTGTTTCTTAATACATT	5333
Db	1333	CAAAATATGATCCGCTCATGAGAAATAAAGCCTCATAAATGCTTTCAAATATATGAAAA	1392
Cp	5332	CAAAATGATGATCCGCTCATGAGAAATAAAGCCTCATAAATGCTTTCAAATATATGAAAA	5373















QY 4665 AGTACAGTAAGTATTTGCGAGTAAATAGTATTTGGGAAACGTTGGTGGATGTCACAGG 4724  
 DB 2294 CATCGTGGTGTACGCTGCTGCTGTTGGTATGCTTTCATTCAGCTCGGTTCCCAACGATC 2353  
 QY 4725 CATCGTGGTGTACGCTGCTGCTGTTGGTATGCTTTCATTCAGCTCGGTTCCCAACGATC 4784  
 DB 2354 AAGCGAGTTATATGATGCTGCTGCTGTTGGTATGCTTTCATTCAGCTCGGTTCCCAACGATC 2413  
 QY 4785 AAGCGAGTTATATGATGCTGCTGCTGTTGGTATGCTTTCATTCAGCTCGGTTCCCAACGATC 4844  
 DB 2414 GATCGTGGTGTACGCTGCTGCTGTTGGTATGCTTTCATTCAGCTCGGTTCCCAACGATC 2473  
 QY 4845 CATCGTGGTGTACGCTGCTGCTGTTGGTATGCTTTCATTCAGCTCGGTTCCCAACGATC 4904  
 DB 2474 TAATTCCTTACTGCTGCTGCTGCTGTTGGTATGCTTTCATTCAGCTCGGTTCCCAACGATC 2533  
 QY 4905 TAATTCCTTACTGCTGCTGCTGCTGTTGGTATGCTTTCATTCAGCTCGGTTCCCAACGATC 4964  
 DB 2534 CAAGTCATTCAGATATGCTGCTGCTGTTGGTATGCTTTCATTCAGCTCGGTTCCCAACGATC 2593  
 QY 4965 CAAGTCATTCAGATATGCTGCTGCTGTTGGTATGCTTTCATTCAGCTCGGTTCCCAACGATC 5024  
 DB 2594 GGATATACGCGGCTGCTGCTGCTGTTGGTATGCTTTCATTCAGCTCGGTTCCCAACGATC 2653  
 QY 5025 GGATATACGCGGCTGCTGCTGCTGTTGGTATGCTTTCATTCAGCTCGGTTCCCAACGATC 5084  
 DB 2654 GGCGGCAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCG 2713  
 QY 5085 GGCGGCAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCG 5144  
 DB 2714 TTAGCGCAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCG 2773  
 QY 5145 TTAGCGCAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCG 5204  
 DB 2774 AGGAAGCAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCG 2833  
 QY 5205 AGGAAGCAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCG 5264  
 DB 2834 ACTCTCTCTTTTCAATATTTAGAGCAATTTATCAGGCTTATGCTGTCATGAGCGGATA 2893  
 QY 5265 ACTCTCTCTTTTCAATATTTAGAGCAATTTATCAGGCTTATGCTGTCATGAGCGGATA 5324  
 DB 2894 CATATTGATGATTTAGAGCAATTTATCAGGCTTATGCTGTCATGAGCGGATA 2953  
 QY 5325 CATATTGATGATTTAGAGCAATTTATCAGGCTTATGCTGTCATGAGCGGATA 5384  
 DB 2954 AGTGGCAC 2961  
 QY 5385 AGTGGCAC 5392

RESULT 13  
 LOCUS SYNBLKSMV 2964 bp DNA circular SYN 04-JUN-1993  
 DEFINITION Bluescribe KS Minus cloning vector.  
 ACCESSION 108784  
 NID 9310728  
 KEYWORDS Synthetic construct DNA.  
 SOURCE  
 ORGANISM  
 TITLE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Unpublished from VecBase 3.0  
 These data and their annotation were supplied to GenBank by Will  
 Gilbert under the auspices of the GenBank Curator Program.  
 Bluescribe KS Minus - Cloning vector  
 ENTRY BLUESKM  
 TITLE Bluescribe KS Minus - Cloning vector  
 DATE 28-JAN-1987  
 #sequence 02-FEB-1987  
 #sequence 04-MAR-1987  
 #sequence 03-APR-1987

ACCESSION VB0077  
 SOURCE artificial  
 REFERENCE  
 #number 1  
 #authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Soroe J.  
 #journal Gene (1987) in press  
 #citation Sequence data from StrataGene  
 #comment sequence correction according to StrataGene comment  
 Obtained from StrataGene on floppy disc.  
 Received 02-FEB-1987 by F. Pfeiffer.  
 1409/10 'AT' to 'TA' to match revised sequence of pBR322  
 Revised 4-MAR-1987 to match sequence of pUC19 on request  
 of StrataGene  
 Polylinker region revised 03-APR-1987 according to StrataGene  
 COMMENT  
 The stand shown corresponds to pUC19c.  
 As in the published sequence of pUC19c, The M13mpl9 lacZ region  
 is on the complementary strand.  
 COMMENT  
 This vector contains the f1 origin so that the minus strand  
 can be obtained upon f1 superinfection.  
 KEYWORDS  
 CROSSREFERENCE  
 #parent  
 VecBase(3):BlueM13m  
 #parent  
 VecBase(3):pUC19, VecSource(3):Prom17, VecSource(3):Prom13,  
 VecSource(3):Bgalks, GenBank(50):PF1  
 #brother  
 VecBase(3):BlueKsp, VecBase(3):BlueSkM  
 PARENT  
 Features of BlueKsm (2964 bp)  
 residue source  
 3- 458 5488-5943 phage f1  
 460- 524 236- 400 pUC19c  
 626- 645 1- 20 T7 promoter  
 653- 760 108- 1 (c) Blues-polylinker  
 772- 791 20- 1 (c) T3 promoter  
 795-2964 448-2617 pUC19c  
 Conflict (cfl) and Mutations (mut): none  
 PARENT  
 Features of BlueKsm (2964 bp)  
 residue source  
 3- 458 5488-5943 phage f1  
 460- 524 236- 400 pUC19c  
 626- 645 1- 20 T7 promoter  
 653- 760 108- 1 (c) Blues-polylinker  
 772- 791 20- 1 (c) T3 promoter  
 795-1031 237- 1 (c) pUC19  
 1032-2964 2686- 754 (c) pUC19  
 Conflict (cfl) and Mutations (mut): none  
 FEATURE  
 643 start of T7-RNA synthesis  
 774 (c) start of T3-RNA synthesis  
 1976-2764 789-1 (c) Ap-R; b-lactamase  
 POLYLINKER  
 KpnI-DraII-ApaI-XhoI-SalI-HindIII-EcoRV-EcoRI-PstI-  
 SmaI-BamHI-SpeI-XbaI-NotI-XmaII-PstXI-SacII-SacI SELECTION  
 #resistance Ap  
 #indicator beta-galactosidase  
 #length 2964 #checksum 1589.  
 SUMMARY  
 Location/Qualifiers  
 1..2964  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 BASE COUNT 708 a 756 c 735 g 755 t  
 ORIGIN  
 Query Match 41.2%; Score 2224; DB 32; Length 2964;  
 Best Local Similarity 99.9%; Pred No 0 006+00;  
 Matches 2226, Conservative 0, Mismatches 2, Indels 0, Gaps 0;



QY 5325 CATATTGGAATGATATTAGAAAAATAAAACAATAGGGTTCCGGCACAATTTCCCGGAAA 5384

Db 2954 AGTGGCAC 2961  
 |||||

QY 5385 AGTGGCAC 5392

RESULT 14

LOCUS U02449 2967 bp DNA circular SYN 29-MAR-1996

DEFINITION Cloning vector pDIRECT, complete sequence.

ACCESSION U02449

KEYWORDS q413815

SOURCE Cloning vector pDIRECT.

ORGANISM Cloning vector pDIRECT.

REFERENCE 1 (bases 1 to 2967)  
 Kitts,P.A.  
 Cloning Vectors On Disc version 1.3  
 UNpublished  
 REFERENCE 2 (bases 1 to 2967)  
 Kitts,P.A.  
 Direct Submission

JOURNAL Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc.,  
 1020 East Meadow Circle, Palo Alto, CA 94303, USA  
 This vector can be obtained from CLONTECH Laboratories, Inc., 1020  
 East Meadow Circle,  
 Palo Alto, CA 94303, USA. To place an order call (415) 424-8222 or  
 (800) 562-2566, extension 1. International customers, please  
 contact your local distributor. For technical information, call  
 (415) 424-8222 or (800) 562-2566, extension 3.  
 This sequence has been compiled from information in the sequence  
 databases, published literature and other sources, together with  
 partial sequences obtained by CLONTECH: this vector has not been  
 completely sequenced. If you suspect there is an error in this  
 sequence, please contact CLONTECH's Technical Service Department at  
 (415) 424-8222 or (800) 562-2566, extension 3 or E-mail  
 TECH@CLONTECH.COM.

FEATURES  
 source  
 1..2967  
 /organism="Cloning vector pDIRECT"  
 /db\_xref="taxon:31799"

BASE COUNT 726 a 756 c 735 g 750 t

ORIGIN

Query Match 41.2%; Score 2224; DB 32, Length 2967;  
 Best Local Similarity 99.9%; Pred. No 0.00e+00;  
 Matches 2226; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 700 GTGGCACTTTTCGGGGAATATGCGCGGAACCCCTATTTCCTTAAATACATT 759  
 |||||

Cp 5392 GTGGCACTTTTCGGGGAATATGCGCGGAACCCCTATTTCCTTAAATACATT 5333  
 |||||

Db 760 CAATATGATATCCGCTCATGAGCAATTAACCCCTGATAAATGCTTCAATATTTGAATA 819  
 |||||

Cp 5332 CAATATGATATCCGCTCATGAGCAATTAACCCCTGATAAATGCTTCAATATTTGAATA 5273  
 |||||

Db 820 GGAAGAGTATGAGTATCAACATTTCCGCTGTCGCCCTTATTCCTTTTTCGGGCAATTT 879  
 |||||

Cp 5272 GGAAGAGTATGAGTATCAACATTTCCGCTGTCGCCCTTATTCCTTTTTCGGGCAATTT 5213  
 |||||

Db 880 GCGTTCTCTGTTTTCGCTCAACCAAGACGCTGTAAGAGTAAAGATGCTGAAGATCAT 939  
 |||||

Cp 5212 GCGTTCTCTGTTTTCGCTCAACCAAGACGCTGTAAGAGTAAAGATGCTGAAGATCAT 5153  
 |||||

Db 940 TGGTGGACAGTGGGTTTACATCAATGATGATGATGATGATGATGATGATGATGATGAT 999  
 |||||

Cp 5152 TGGTGGACAGTGGGTTTACATCAATGATGATGATGATGATGATGATGATGATGATGAT 5093  
 |||||

Db 1000 TTGGCCCGGGAAGAGGTTTTCCTTCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCG 1059  
 |||||

Cp 5092 TTGGCCCGGGAAGAGGTTTTCCTTCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCG 5033  
 |||||

Db 1060 TATTATCCGCTATTGACGCGGCGCAAGAGTAACCTCGGTCGCGGCAATACATATATTCAGSA 1119  
 |||||

Cp 5032 TATTATCCGCTATTGACGCGGCGCAAGAGTAACCTCGGTCGCGGCAATATATTCAGSA 4973  
 |||||

Db 1120 ATGACTTGGTTGAGTACTCACCAGTACAGAAAAAGCATCTTTACGGATGGCATGACAGTAA 1179  
 |||||

Cp 4972 ATGACTTGGTTGAGTACTCACCAGTACAGAAAAAGCATCTTTACGGATGGCATGACAGTAA 4913  
 |||||

Db 1180 GAGAAATATGCACTGCTGTCATATCCATATCCATATGATATGATATGATATGATATGATAT 1239  
 |||||

Cp 4912 GAGAAATATGCACTGCTGTCATATCCATATGATATGATATGATATGATATGATATGATAT 4853  
 |||||

Db 1240 CAACGATCGGAGCACGACGAGGAGCTAACCGCTTTTTCGCAACATGCGGCGATCATGATA 1296  
 |||||

Cp 4852 CAACGATCGGAGCACGACGAGGAGCTAACCGCTTTTTCGCAACATGCGGCGATCATGATA 4793  
 |||||

Db 1300 CTCGCTTGTATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCGTGACA 1359  
 |||||

Cp 4792 CTCGCTTGTATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCGTGACA 4733  
 |||||

Db 1360 CCAGCATGCGCTGTAGTAATGGCAACACGCTGCGCAACATTTAACTGCGCAACTACTTA 1419  
 |||||

Cp 4732 CCAGCATGCGCTGTAGTAATGGCAACACGCTGCGCAACATTTAACTGCGCAACTACTTA 4673  
 |||||

Db 1420 CTCAGCTTCCCGGCAACAAATTAATAGACTGGATGGAGCGGATAAAGTTGCGAGGACAC 1479  
 |||||

Cp 4672 CTCAGCTTCCCGGCAACAAATTAATAGACTGGATGGAGCGGATAAAGTTGCGAGGACAC 4613  
 |||||

Db 1480 TTCTGGCTCGGCCCTTCGGCTGGCTGGTTTATTTATTTATTTATTTATTTATTTATTTATTT 1539  
 |||||

Cp 4612 TTCTGGCTCGGCCCTTCGGCTGGCTGGTTTATTTATTTATTTATTTATTTATTTATTTATTT 4553  
 |||||

Db 1540 GTGGCTTCGCGGTATCATTTGAGCACTGGGCGCACATGATGATGATGATGATGATGATGATG 1599  
 |||||

Cp 4552 GTGGCTTCGCGGTATCATTTGAGCACTGGGCGCACATGATGATGATGATGATGATGATGATG 4493  
 |||||

Db 1600 TTATCTACAGCAGCGGAGTCAAGCAACTATGATGATGATGATGATGATGATGATGATGATG 1659  
 |||||

Cp 4492 TTATCTACAGCAGCGGAGTCAAGCAACTATGATGATGATGATGATGATGATGATGATGATG 4433  
 |||||

Db 1660 TAGTGGCTCACTGATTAAGCATTTGTAAGTGTGAGCAACCAAGTTTACTCATATATACTTT 1719  
 |||||

Cp 4432 TAGTGGCTCACTGATTAAGCATTTGTAAGTGTGAGCAACCAAGTTTACTCATATATACTTT 4373  
 |||||

Db 1720 AGATTGATTTAAAACTTCATTTTAAATTTAAAGGATCTAGTGAAGATCCCTTTTIGATA 1779  
 |||||

Cp 4372 AGATTGATTTAAAACTTCATTTTAAATTTAAAGGATCTAGTGAAGATCCCTTTTIGATA 4313  
 |||||

Db 1780 ATCTCATGACCAAAATCCCTTAAAGTGTGATTTTCGTTTCCACTGAGCGTCAGACGCGGTAG 1839  
 |||||

Cp 4312 ATCTCATGACCAAAATCCCTTAAAGTGTGATTTTCGTTTCCACTGAGCGTCAGACGCGGTAG 4253  
 |||||

Db 1840 AAAAGATCAAGAGTCTCTTIGAGATCCCTTTTTCGCGGCTAACTCTCTCTCTCTCTCTCTCT 1899  
 |||||

Cp 4252 AAAAGATCAAGAGTCTCTTIGAGATCCCTTTTTCGCGGCTAACTCTCTCTCTCTCTCTCTCT 4193  
 |||||

Db 1900 CAAAAAACCCACCGCTACCGAGCGGTGTTTGTTCGCGGATCAAGAGTACCAACTCTTT 1959  
 |||||

Cp 4192 CAAAAAACCCACCGCTACCGAGCGGTGTTTGTTCGCGGATCAAGAGTACCAACTCTTT 4133  
 |||||

Db 1960 TTCGGAAGTAACTGGCTTCAGCAGAGCGCATACCAAAACGTCGCTCTCTCTCTCTCTCTCT 2019  
 |||||

Cp 4132 TTCGGAAGTAACTGGCTTCAGCAGAGCGCATACCAAAACGTCGCTCTCTCTCTCTCTCTCT 4073  
 |||||

Db 2020 CGTAGTTAGCGCACCACTTCAAGAACTCTAGCACCGCTACATACCTCGCTCTCTCTCTCT 2079  
 |||||

Cp 4072 CGTAGTTAGCGCACCACTTCAAGAACTCTAGCACCGCTACATACCTCGCTCTCTCTCTCT 4013  
 |||||

Db 2080 TCCTGTTTACCAGTGGCTGCTGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2139  
 |||||

Cp 4012 TCCTGTTTACCAGTGGCTGCTGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3953  
 |||||

Db 2140 GACGATAGTTACCGGATAGAGCGGCGGCGGTTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 2199  
 |||||

[illegible]

QY 4071 CGGTACACATAGAGGACAGATTTGGTATCTGGCTCTGCTGAAGCCAGTTACCTTCGG 4130  
 Db 6213 AAAAAGAGTTGGTAGCTCTTATCCGCGCAACAAACAGCGCTGGTAGCGGTGTTTTT 6272  
 QY 4131 AAAAAGAGTTGGTAGCTCTTATCCGCGCAACAAACAGCGCTGGTAGCGGTGTTTTT 4190  
 Db 6273 TGTTCGAAGCAGCAGATTAACGCGCAGAAAAAGAGATCTCAAGAGATGCTTCATGCTT 5332  
 QY 4191 TGTTCGAAGCAGCAGATTAACGCGCAGAAAAAGAGATCTCAAGAGATGCTTCATGCTT 4250  
 Db 6333 TTCTACGGGCTCTGACGCTCAGTGGAGCAAACTCAGGTTAAGGATTTTGGTCATGAG 6392  
 QY 4251 TTCTACGGGCTCTGACGCTCAGTGGAGCAAACTCAGGTTAAGGATTTTGGTCATGAG 4310  
 Db 6393 ATTATCAAAAGAGATCTTCACCTAGATGCTTTTAAATTAATAATGAAGTTTAAATCAAT 6452  
 QY 4311 ATTATCAAAAGAGATCTTCACCTAGATGCTTTTAAATTAATAATGAAGTTTAAATCAAT 4370  
 Db 6453 CTAAGTATATAGTAAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGCACC 5512  
 QY 4371 CTAAGTATATAGTAAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGCACC 4430  
 Db 6513 TATCTACGCGATCTGCTATTTGCTTCATCCATAGTTGCTGACTCCCGCTGCTGATGAT 6572  
 QY 4431 TATCTACGCGATCTGCTATTTGCTTCATCCATAGTTGCTGACTCCCGCTGCTGATGAT 4490  
 Db 6573 AACTACGATACGGGAGGCTTACCATCTGGCCCGAGTGTGCAATGATACCGCGAGACCC 6632  
 QY 4491 AACTACGATACGGGAGGCTTACCATCTGGCCCGAGTGTGCAATGATACCGCGAGACCC 4550  
 Db 6633 ACGTCCAGGCTCCAGATTTATCAGCAATTAACAGCAGCGGAGCGGCGAGCGAG 6692  
 QY 4551 ACGTCCAGGCTCCAGATTTATCAGCAATTAACAGCAGCGGAGCGGAGCGAGCGAG 4610  
 Db 6693 AAGTGGTCTGCACTTTATCCGCTCCAGTCTATTAATGTTGCGGGAAGCTAG 6752  
 QY 4611 AAGTGGTCTGCACTTTATCCGCTCCAGTCTATTAATGTTGCGGGAAGCTAG 4670  
 Db 6753 AGTAAGTAGTTCGCGAGTTAATAGTTGCGCAACGTTGTTGCCATTCGCTACAGGATGCT 6812  
 QY 4671 AGTAAGTAGTTCGCGAGTTAATAGTTGCGCAACGTTGTTGCCATTCGCTACAGGATGCT 4730  
 Db 6813 GGTGTCAGGCTCGTGTGTTGGTATGCTTCATTCAGCTCCGTTTCCCAACGATCAAGCGG 6872  
 QY 4731 GGTGTCAGGCTCGTGTGTTGGTATGCTTCATTCAGCTCCGTTTCCCAACGATCAAGCGG 4790  
 Db 6873 ACTTACATGATCCCGCATGTTGTGCAAAAGCGGTTAGTCTCTCGTCCCTCCGATGCT 6932  
 QY 4791 ACTTACATGATCCCGCATGTTGTGCAAAAGCGGTTAGTCTCTCGTCCCTCCGATGCT 4850  
 Db 6933 TGTGAGAAGTAAAGTTGGCCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATTC 6992  
 QY 4851 TGTGAGAAGTAAAGTTGGCCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATTC 4910  
 Db 6993 TCTTACTGTGATGCTATCCGCTAGATGCTTTTCTGCTGATGCTGAGTACTCAACCAAGTC 7052  
 QY 4911 TCTTACTGTGATGCTATCCGCTAGATGCTTTTCTGCTGATGCTGAGTACTCAACCAAGTC 4970  
 Db 7053 ATTCTGAGAATAGTGTATGGCGGACCGAGTTGCTCTTGGCCGCGCTCAATACGGGATAA 7112  
 QY 4971 ATTCTGAGAATAGTGTATGGCGGACCGAGTTGCTCTTGGCCGCGCTCAATACGGGATAA 5030  
 Db 7113 TACCGCGGCAATAGCAGAACTTTAAAGTGCCTCATCATGTTGAAAAAGCTTCTTCGGGGG 7172  
 QY 5031 TACCGCGGCAATAGCAGAACTTTAAAGTGCCTCATCATGTTGAAAAAGCTTCTTCGGGGG 5090  
 Db 7173 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTCCGATTAACCCACTCGTGACC 7232  
 QY 5091 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTCCGATTAACCCACTCGTGACC 5150  
 Db 7233 CAACTGATCTTCAGCATCTTTTACTTTTCAAGCGGTTTCTGGTGAGCAAAAAACAGGAAG 7292  
 QY 5151 CAACTGATCTTCAGCATCTTTTACTTTTCAAGCGGTTTCTGGTGAGCAAAAAACAGGAAG 5210

Db 7293 GCAAAATGCGCAAAAAAGGGAATAGCGCGACACGGAATGTTGAATACTCATACTCTT 7352  
 QY 5211 GCAAAATGCGCGCAAAAAAGGGAATAGCGCGACACGGAATGTTGAATACTCATACTCTT 5270  
 Db 7353 CCTTTTCAATATATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 7412  
 QY 5271 CCTTTTCAATATATTGAAGCAATTTATCAGGGTTATTGTCTCATGAGCGGATACATATT 5330  
 Db 7413 TGAATGATTTAGAAAAATTAACAAATAGGGGTTCCGGGCACTTTTCCGCAAAAGTCC 7472  
 QY 5331 TGAATGATTTAGAAAAATTAACAAATAGGGGTTCCGGGCACTTTTCCGCAAAAGTCC 5390  
 Db 7473 AC 7474  
 QY 5391 AC 5392

Search completed: Mon Mar 15 23:40:40 1999  
 Job time : 13760 secs.





\*\*\*\*\*  
M O S E R E H  
\*\*\*\*\*  
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsearch\_n n.a. n.a database search using Smith-Waterman algorithm  
Run on Tue Mar 15 02:50:01 1999. Maspar time 759.81 seconds  
965.496 Million cell updates/sec  
Tabular output not generated

Title: >US-09-020-716-6  
Description: (1-5392) from US09020716 seq  
Perfect Score: 5392  
N.A. Sequence: 1 CTAAATTGTAAAGCGTTAATA ATTTCGCCGAAAAGTGCAC 5392  
Comp: GATTTAACATTGCAATAT TAAAGGCGCTTTTCACGGTG

Scoring table: TABLE default  
Gap 5

Nmatch STD : DBase 0: Query 0

Searched: 189442 seqs, 58025449 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-genes32  
1:part1 2:part3 3:part4 4:part5 5:part6 6:part7 7:part8  
8:part9 9:part10 10:part11 11:part12 12:part13 13:part14  
14:part15 15:part16 16:part17 17:part18 18:part19  
19:part20 20:part21 21:part22 22:part23 23:part24  
24:part25 25:part26 26:part27 27:part28 28:part29  
29:part30 30:part31 31:part32 32:part33 33:part34  
34:part35 35:part36 36:part37 37:part38 38:part39  
39:part40 40:part41

Statistics: Mean 10.519; Variance 7.117; scale 1.478  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES							
Result	No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2222	41.2	5178	27	T49876	pTet-Splice.	0.00e+00
2	2220	41.2	6206	27	T49877	Autoregulatory vector	0.00e+00
3	2222	41.2	7474	31	T71320	Plasmid pPB73 encodin	0.00e+00
4	2203	40.8	5314	2	Q19576	Plasmid pMTV1 contain	0.00e+00
5	2201	40.8	3681	2	Q19578	Plasmid pKSE15.	0.00e+00
6	2198	40.8	3699	39	V14340	Plasmid pBSGFP expres	0.00e+00
7	2202	40.8	4145	7	Q40279	Sequence of clone pS2	0.00e+00
8	2198	40.8	4277	7	Q40280	Sequence of clone pS2	0.00e+00
9	2200	40.8	4732	15	O84596	Plasmid GS control. N.	0.00e+00
10	2200	40.8	5042	15	O84594	Plasmid glucoamylase	0.00e+00
11	2202	40.8	5356	25	T43794	Plasmid pRIPHAT (rat	0.00e+00
12	2202	40.8	5534	25	T43137	PUMIGIT sequence incl	0.00e+00
13	2199	40.8	5965	7	Q40291	Sequence of clone pP2	0.00e+00

C 14 2198 40.8 7287 39 V02042  
15 2196 40.7 4539 14 O87347  
16 2196 40.7 10930 20 O81236  
17 2196 40.7 10930 20 O81236  
18 2196 40.7 10930 20 O81236  
19 2196 40.7 10930 20 O81236  
20 2196 40.7 10930 20 O81236  
21 2196 40.7 10930 20 O81236  
22 2196 40.7 10930 20 O81236  
23 2196 40.7 10930 20 O81236  
24 2196 40.7 10930 20 O81236  
25 2196 40.7 10930 20 O81236  
26 2196 40.7 10930 20 O81236  
27 2196 40.7 10930 20 O81236  
28 2196 40.7 10930 20 O81236  
29 2196 40.7 10930 20 O81236  
30 2196 40.7 10930 20 O81236  
31 2196 40.7 10930 20 O81236  
32 2196 40.7 10930 20 O81236  
33 2196 40.7 10930 20 O81236  
34 2196 40.7 10930 20 O81236  
35 2196 40.7 10930 20 O81236  
36 2196 40.7 10930 20 O81236  
37 2196 40.7 10930 20 O81236  
38 2196 40.7 10930 20 O81236  
39 2196 40.7 10930 20 O81236  
40 2196 40.7 10930 20 O81236  
41 2196 40.7 10930 20 O81236  
42 2196 40.7 10930 20 O81236  
43 2196 40.7 10930 20 O81236  
44 2196 40.7 10930 20 O81236  
45 2196 40.7 10930 20 O81236

ALIGNMENTS

RESULT 1  
ID T49876 standard: DNA: 5178 BP.  
AC T49876:  
DE pTet-Splice.  
KW pTet-Splice; pTet-UTR; tetracycline transactivator, promoter;  
KW gene expression; transgenic animal; animal model; drug screening;  
KW vector, ds.  
FH Synthetic.  
FS Key  
FT Location/Qualifiers  
FT misc\_rna  
FT complement (2591)  
FT /\*tag= a  
FT /\*note= putative start site of trxn  
FT complement (2516..2522)  
FT /\*tag= b  
FT tata\_signal  
FT WO9604946-A1.  
PN 19-DEC-1996.  
PD 07-JUN-1996: U010109  
PF 07-JUN-1996: US-474169.  
PR (UYIA ) UNIV YALE.  
PI Schatz DG:  
DR WPI: 97-077273/07.  
PT Nucleic acid encoding tetracycline transactivator fusion protein -  
PT provides rapid and reversible control of gene expression, e.g. for  
PT creating animal models for drug screening  
PS Example 1, Fig 9b-g; 82pp; English.  
CC Vector pTet-Splice (T49876) is utilized in the construction of  
CC autoregulatory vector plasmid pTet-trak (T49877). Splice-PA was  
CC made by ligating the SV40 small t antigen intervening sequence and  
CC the SV40 early polyA sequence into pBSK1+. The XhoI-SalI fragment  
CC of pUHI3.3 (contg. 7 copies of the tet operator upstream of minimal  
CC promoter tetp) was cloned upstream of the splice/polyA sequence of  
CC pSplice-PA to form pTet-Splice. A modified tetracycline  
CC transactivator (tTA) gene was cloned into pTet-Splice to form pTet-  
CC trak. The construct provides rapid, reversible control of gene  
CC expression in eukaryotic cells of transgenic animals, e.g. for  
CC heterologous protein, RNA or antisense sequence production, or for







Db 5853 tttccgacccctgcgcttaccggatatacctgtccgcccctttctcccttcgggaagcggtggc 5912  
 QY |||||  
 Db 3771 TGTTCGAGCCCTGCCGCTTACCGGATACCTGTCCGCCCTTCCTCCCTCGGGAAGCGTGGC 3830  
 QY |||||  
 Db 5913 gctttctatagctacgctgtagtattatcagttcgatgtagtgcgttcctccaagct 5972  
 QY |||||  
 Db 3831 GCTTTCTCATAGCTCAGCGTGTAGTATATCTCATGTTGGTGTAGTGTGCTGCTCCAAAGCT 3890  
 QY |||||  
 Db 5973 ggcctgtgacgaagaccccccgttcagccagccgctgcgccccttatccggaataactatcg 6032  
 QY |||||  
 Db 3891 GGGCTCTGTGTCAGCAACCCGCCGCTACAGCCCGGACCGCTTCGCCCTTATCCGGAATACTATCG 3950  
 QY |||||  
 Db 6033 tcttgagtcacaccccggttaagacacgacttatccgacactggcagcagccactggttaacag 6092  
 QY |||||  
 Db 3951 TCTTGAGTCCCAACCCGCTTAAGACACGACTTATCCGACCTGCGAGCAGCACTGGTAACAG 4010  
 QY |||||  
 Db 5093 gattacagagcagagtagttagggcggtgctacagagttcttgaagtggtggcctaacta 6152  
 QY |||||  
 Db 4011 GATTACAGAGCGAGGTATGTAGGGCGTGTCTACAGAGTCTTTGAAGTGTGGCCTAACCTA 4070  
 QY |||||  
 Db 6153 cggctacactagaagacagtagtttggtagctgctgctgctgctgagcagttaccttcgg 6212  
 QY |||||  
 Db 4071 CGGCTACACTAGAACGACAGTATTTGGTATCTCGGCTCTGCTGAAGCCAGTTACCTTCGG 4130  
 QY |||||  
 Db 6213 aaaaagagttggttagctcttggtagcgggcaaaaacaccccgctggtgtagcgtgtttttt 6272  
 QY |||||  
 Db 4131 AAAAAGAGTTGTTAGTCTTGTATCCGGCAACCAACACCGCTGGTAGCGGTGTTTTT 4190  
 QY |||||  
 Db 6273 tgtttcgaagcaacatatacgcgcagaaaaaaagagatctcaagaagatcctttgatctt 6332  
 QY |||||  
 Db 4191 TGTTCGAGCAGCAGTATACGCGCAGAAAAAAGGATCTCAAGAAAGATCCTTTGATCTT 4250  
 QY |||||  
 Db 6333 tctcaggggtctgcagctcagtgaaacgaacactcaaggttaagggattttgggtcatgag 6392  
 QY |||||  
 Db 4251 TTTACGGGCTCTGAGCTTCAGTGGAAACCAACTCAGTTAAGGGATTTTGGTCATGAG 4310  
 QY |||||  
 Db 6393 attatcaaaaagagcttccacttagatccttttaattaaaatgaagtttaataatcaat 6452  
 QY |||||  
 Db 4311 ATTATCAAAAAGAGATCTTACCTAGTAGTCTTTTAAATTAATAATGAAGTTTAAATCAAT 4370  
 QY |||||  
 Db 6453 cttaagtatatagtaaaacttgctcagagttacaaatgcttaatacagtaggagacc 6512  
 QY |||||  
 Db 4371 CTAAAGTATATAGTAATAACTTGGTCTGACAGTTTACCAGTCTTAATCAGTAGGACAC 4430  
 QY |||||  
 Db 6513 tatctcagcagctgtctattttgcttcacatagttgctgactcccccgctgtagat 6572  
 QY |||||  
 Db 4431 TATCTCAGCGATCTGTCTATTTCGTTTATCCATAGTTGCTTACCTCCCGTCTGTTAGAT 4490  
 QY |||||  
 Db 6573 aactacagatcagggaggttaccatctgccccagtgctgcaatgtatccgcagagacc 6632  
 QY |||||  
 Db 4491 AACTACGATACGGGAGGGCTTACCATCTGGCCCGCAGTGTGCAATGATACCGCGAGACCC 4550  
 QY |||||  
 Db 6633 acgctacccggtccagatttatcagcaataaaccagcagccggaagggcgagcgagcag 6692  
 QY |||||  
 Db 4551 ACGCTACCCGGCTCCAGATTTATCAGCAATTAACACGCGCGGGAAGGGCGAGCGAG 4610  
 QY |||||  
 Db 6693 aatggtctctgcaactttatccgctccatccagctctatattattgttgcgggaagctag 6752  
 QY |||||  
 Db 4611 AAGTGTCTCTGCAACTTTATCCGCCCTCCATCCAGTCTATTAATTTGTGCCGGAGCTAG 4670  
 QY |||||  
 Db 6753 agtaagtagtccaggttaagtttggcgaagctgttgcattgtacagagcattcgt 6812  
 QY |||||  
 Db 4671 AGTAAGTAGTTCGCCAGTTAATAGTTTGGCAACCGTGTGTCATTTGTACAGCATCGT 4730  
 QY |||||  
 Db 6813 ggtgtcacgctcgtctgttggtagcttcattcagctccggttcccccaactcaagcgc 6872  
 QY |||||  
 Db 4731 GGTGTCACTGCTGCTGTTGGTAGTGTGCTTCACTCAGCTCCGTTCCCAACGATCAAGCGC 4790  
 QY |||||  
 Db 6873 agttacatgacccccctgttgtgcaaaaaagcggttagctccttcggtccctccgtagct 6932  
 QY |||||  
 Db 4791 AGTTACATGATCCCCCATGTTGTGGCAAAAAGCGGTTAGCTCCTTCGTTCCGATCGT 4850  
 QY |||||

Db 6933 tdtcagaagtaagtggccgcagtggttatcactcatgttatggcagcactgcataatc 6992  
 QY |||||  
 Db 4851 TGTCAAGAGTAAGTTGGCCGAGTGTATCACTCATGTTATGTTATGTCAGCACTGCAATATC 4910  
 QY |||||  
 Db 6993 tcttactgtatgcactccgtaaatgcttttctgtgactggtgagtagtcaaccgaatc 7052  
 QY |||||  
 Db 4911 TCTTACTGTATGTCATCCGTAAGAAGTCTTCTGTACTGTGAGTACTCAACCAAGTC 4970  
 QY |||||  
 Db 7053 attctgaaatagttatgctgcgcagcagagttgcttcttgcggcgtcaatacaggaataa 7112  
 QY |||||  
 Db 4971 ATTCTGAGAAATAGTGTATGCGCGACCGAGTTGCTCTTGGCCGCGTCAATACCGGATAA 5030  
 QY |||||  
 Db 7113 taccgcgcacatagcagaactttaaaagtgctcatcatttggaacacggttcttcggugcg 7172  
 QY |||||  
 Db 5031 TACCGCGCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAAACTGTTCTTGGGAGC 5090  
 QY |||||  
 Db 7173 aaactctcaaggtatcttaccgctggttgagatccagttcagatgaacccactcgtgcacc 7232  
 QY |||||  
 Db 5091 AAAACTCTCAAGGATCTTTACGCTGTTGAGATCCAGTTCGATGTAACCCACTCTGTGACC 5150  
 QY |||||  
 Db 7233 caactgactctcagcatctttactttccacagcgtttctggtggaacaaaacagaaag 7292  
 QY |||||  
 Db 5151 CAACTGATCTTCAGCATCTTTTACTTTCACGAGGTTTCTTGGGTGAGCAAAAACAGGAAG 5210  
 QY |||||  
 Db 7293 gcaaatccgcgcacaaagggaataaaggcgacacacggaatgttgatactactacactt 7352  
 QY |||||  
 Db 5211 GCAAAATGCGCAAAAAGGAAATAAGGCGCACACGGAATGTTGAATACTACTACTCTT 5270  
 QY |||||  
 Db 7353 ccttttcaatattattgaagcatttatcagggttattgtctctcagcagcagatacatt 7412  
 QY |||||  
 Db 5271 CTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAIT 5330  
 QY |||||  
 Db 7413 tgaatgtatttagaaaaataaacaataaggggtccgcgacatttcccccgaagtgcc 7472  
 QY |||||  
 Db 5331 TCAATGTATTAGAAAAATAACAATAAGGGGTTCCCGGCACATTTCCCGGAAAAAGTGC 5390  
 QY |||||  
 Db 7473 ac 7474  
 QY ||  
 Db 5391 AC 5392

## RESULT 4

ID Q13576 standard; DNA: 5314 BP.

AC Q13576;

DT 02-DEC-1991 (first entry)

DE plasmid pMTV1 containing 3 multiple cloning sites.

KW recombinant expression plasmid; phage phi X174; protein E;

KW bacterial ghost; ss.

OS Synthetic.

PN WO9113155-A.

PD 05-SEP-1991.

PF 19-FEB-1991; E00308.

PR 24-FEB-1990; DE-005874.

PA (BOEF), BOEHRINGER MANNHEIM GMBH.

PI Lubitz W, Szostak MP;

WP1: 91-281471/38.

PT Carrier recombinant protein for vaccines against HIV, etc.

PT obt'd. by expressing fusion protein gene in gram negative bacteria

PT and gene that encodes for lytic membrane protein

PS Example 10; Page 28-29; 45pp; German.

CC The carrier plasmid pMTV1 comprises the phage phi X174 E protein

CC gene, the ampicillin resistance gene, the lambda ci857 repressor

CC gene and P(R) from the lambda promoter/operator system. There are

CC also lac operon sequences and three multiple cloning sites. A

CC foreign gene can be inserted into the second multiple cloning site

CC (mcs2). The protein encoded by the foreign gene is especially a

CC viral antigenic peptide, e.g. of HIV, HBV or EBV.

CC See also Q13577 and Q13578.

SQ Sequence 5314 BP; 1438 A; 1276 C; 1310 G; 1290 T;

Query Match 40.9%; Score 2303; DB 2; Length 5314;

Best Local Similarity 99.5%; Pred. No. 0.00e+00;

Matches 2215; Conservative 0; Mismatches 12; Indels 0; Gaps 0;



```

QY 5326 ATATTGTAATGATTTAGAAAAATAAATAAGGGTTCCGGGACACATTCCCGCAAAA 5385
Db 5305 gtacac 5311
QY 5386 GTGGCAC 5392

RESULT 5
ID Q13578 standard; DNA; 3681 BP.
AC Q13578:
DT 02-DEC-1991 (first entry)
DE Plasmid pKSEL5
KW recombinant expression plasmid; phage phi X174; protein E;
KW bacterial host; ss.
OS Synthetic
FN W9114155-A.
PD 05-SEP-1991.
PE 19-FEB-1991.
PE 24-FEB-1990. PE-005874
PA (BOE) BOEHRINGER MANNHEIM GMBH.
PI Lubitz W. Szostak MP.
DR WPI: 91-281471/38.
PT Carrier recombinant protein for vaccines against HIV, etc.
PT abd. by expressing fusion protein gene in gram negative bacteria
PT and gene that encodes for lytic membrane protein
PS Example 1: Page 33-34; 45pp; German.
CC Plasmid pKSEL5 comprises three multiple cloning sites (mcs1-3), the
CC ampicillin resistance gene, regions of the lac operon and a
CC partial phage phi X174 E' target sequence. This plasmid can be used
CC as a carrier vector according to the invention for the insertion of
CC sequences encoding viral antigenic peptides, e.g. the HIV gp41
CC sequence was isolated as a HindIII/PvuII fragment from plasmid
CC pPH14. The fragment also included a linker sequence and the last 45
CC codons of gp120. The fragment was ligated to pKSEL5 which had been
CC cleaved with AccI and end-filled. Alternatively, a new carrier
CC vector, denoted pMV1 (see p13576), can be constructed from pKSEL5
CC which has been partially digested with SspI. It is ligated to a
CC SspI-braff fragment from pMD1 (see Q13577). Foreign sequences are
CC cloned into mcs2.
SU Sequence 3681 BP: 949 A: 917 C: 932 G: 883 T:

Query Match 40.8% Score 2201: DB 2: Length 3681:
Best Local Similarity 99.4% Pred No 0.00e+00:
Matches 2214: Conservative 0: Mismatches 13: Indels 0: Gaps 0:

Db 1452 tcgacctcgaggggggcccggtaccagcttttgccttttagtgagggttaattccg 1511
QY 3156 TCGACCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3225
Db 1512 agcttgacataatcatgctatgctatgcttttccctgtgaaattgtatccgctcacaatt 1571
QY 3226 CGCTTGCGGTAATCATGCTCATAGCTGTTTCCGTGTGAAATGTTATCCGCTCAAAAT 3285
Db 1572 ccacacacatagagcgaagacataaagtataagcctgggtgctcctaaagagtggg 1631
QY 3286 CCACACAAACATACAGCGCGGAGACATAAAGTGTAAGAGCGTGGGGTCCCTAATGAGTGAGC 3345
Db 1632 taactcacattaattgctgtgctgctactgcccgcctttccagtcgggaaacctgtcgtgc 1691
QY 3446 TAACATCAATTAATTGCTGTGCGCTACTGCCCGCTTCCAGTCGGAAACCTGTCGTGC 3405
Db 1592 caagctgattaatgaatcgcgaacgcggcgaggagagcgatttgcgtattggcgctct 1751
QY 3496 CAGCTGCAATTAATGAATCGGCCAACCGCGGGGAGAGGGGTTTGGCTATTGGCGCTCT 3465
Db 1752 tccacttccctgcctcactgactgcctgcctcactgctgctgctgctgctgctgctgct 1811
QY 3456 TCCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3525
Db 1812 gctcactcaagggcggttaatcacgtttatccacagaatcaggggataacacagggaagaac 1871
QY 3526 GTTACTCAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1895

```

```

Db 1872 atgtgagcaaaagccagcaaaagccagcaaaagccagcaaaagccagcaaaagccagc 1931
QY 3586 ATGTGACCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAG 3645
Db 1932 ttccataggctcgcccccctcagcagcagcagcagcagcagcagcagcagcagcagcag 1991
QY 3646 TTCCATAGGCTCGCCCCCTCAGCAGCAGCAGCAAAATCGACGCTCAAGTCAGACGCTGG 3705
Db 1992 cgaaccccaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2051
QY 3706 CGAAACCCGACAGCAGCTATAAGATACCAAGGCTTCGCAAGCTTCGCAAGCTTCG 3765
Db 2052 tctcctgttccgacctgcccgttaaccqatacctctcgcgccttctcctcctcgagaac 2111
QY 3766 TCCTCTGTTCGGACCTCGCGGTTCACCGGATACCTTCGCGGTTCGCGGTTCGCGG 2225
Db 2112 gtggcgcttctcaatgctcagcagcagcagcagcagcagcagcagcagcagcagcag 2171
QY 3826 GTGGCGCTTTCATAGCTCAGCTGTAGGATATCTCAGTTCGCTGTAGGCTGTAGGCT 2285
Db 2172 aagctaggcgtgctgcagcaaaccccccgttcagcagcagcagcagcagcagcagcag 2231
QY 3886 AAGCTGGGCTTGTGACGAACCCCGCTTCAGCGGAGAGCTTCGCGGTTCGCGGTTC 3945
Db 2232 tctcgttctgagtcacacaccccggttaacacacagcagcagcagcagcagcagcagc 2291
QY 3946 TATCGCTTTCAGTTCACACCGCGGTAAACACACAGCAGCTTATTCGCTACGTAACAC 4005
Db 2292 aacagcattagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2351
QY 4006 AACAGATAGCAAGAGCAGGATGTAGGCTGTAGGCTGTAGGCTGTAGGCTGTAGGCT 4065
Db 2352 aactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2411
QY 4066 AACTACGCTTACACTAGAGGAGCAGCTTATTCGCTACGCTACGCTACGCTACGCTAC 4125
Db 2412 ttccgaaaaagcgttgcgtgctgctgctgctgctgctgctgctgctgctgctgctgct 2471
QY 4126 TTCGAAAAAGAGCTTGGTATGATTCGCTACGCTACGCTACGCTACGCTACGCTACG 4185
Db 2472 ttcttctgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2531
QY 4186 TTTTGTGTGCAAGAGCAGCTATTCAGCTACGCTACGCTACGCTACGCTACGCTACGCT 4245
Db 2532 atcttctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2591
QY 4246 ATCTTTTCTACGCGGCTCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTAC 4305
Db 2592 atdagattatcaaaaaaggatcttccactagatccttttaaatataaaatgaagttttaa 2651
QY 4306 ATGAGATTATCAAAAAGGATCTTCAGCTACGCTACGCTACGCTACGCTACGCTACG 4365
Db 2652 tcaatcaaaagtatataatgataaaacttactctgacagcttaccatgcttactgctgtaa 2711
QY 4366 TCAATCTAAAGTATATATGATGAAGCTTGGCTGACAGCTTACCAATGCTTAACTACAG 4425
Db 2712 gacactctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2771
QY 4426 GCACCTATCTCAGCGATCTCTTATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4485
Db 2772 tagataactacgatacagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2831
QY 4486 TAGATAACTACGATACGGGAGGCTTACCATCTGCCCCAGTGTCTGCAATGATACCGG 4545
Db 2832 gacccagcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2891
QY 4546 GACCCAGCGCTACCGGCTCCAGATTTATCAGCAATTAACACAGCTACGCTACGCTAC 4605
Db 2892 cgcagaagtggctcctcaacttaccgctcctcctcctcctcctcctcctcctcctcct 2951
QY 4606 CGCAGAAAGTGTCTCTCAACTTTATTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4665

```







```

QY 3911 CCGTTCAGCGAGCGGCTGGCCCTTATCCGCTAACTATCGCTTCAGTCCACCCGGTAA 3970
DB 2275 gacacgacttatcgccactgcacagccactggtaacagagattagcagagcagatagc 3934
QY 3971 GACACGACTTATCGCCACTGCGCAGCAGCCACTGGTACAGGATTAGCAGAGGAGTATG 4030
DB 2335 tagcggtgctacagagttcttgaagtgggtggcctaactacggtcacactagagagacag 2394
QY 4031 TAGGGGCTGCTACAGAGTCTTGAAGTGGTGGCTTAAGTACAGGCTACACTAGAGGACAG 4090
DB 2395 tatttgatctgagctctgctgagccaggttaccttcggaaaaagattggtagctctt 2454
QY 4091 TATTGGTATGTGGCTCTGTGTGAAGCAGATTACCTTCGGGAAAAAGAGTTGGTAGCTCTT 4150
DB 2455 gatacgcgaacaaacacccactgtagcgggtgttttttggttgtaacacacagatca 2514
QY 4151 GATCCGGCAACCAACTACCGCTGTGAGCGTGTGTTTTTGTGGTGGTGGTGGTGGTGGT 4210
DB 2515 cgcgcagaaaaaagagatctcaagaadactcttttgatctcttttcaagggctgcagcctc 2574
QY 4211 CAGTGAAGAAAAAGGATCTCAAGAGATGCTTTGATCTTTTCTACGGGGTCTGACGCTC 4270
DB 2575 atggaacaaacactcaacttaagagattttgttcacagattatcaaaaagatcttca 2634
QY 4271 AGTGAAGCAAAACCTACGCTTAAAGGATTTTGGTCAATGAGATTATCAAAAAGGATCTTCA 4330
DB 2635 cctgagctctttaaataaaatgaagttttaaatacaatcaatcaatcaatcaatcaatcaat 2694
QY 4331 CCGATGCTTTTAAATTAATAATGAAGTTTTTAATCAATCAATCAATCAATCAATCAAT 4390
DB 2695 cttggtctgacagttaccactgcttaatacagtgagggcacctatctcagcagatctctat 2754
QY 4391 CTGTGCTCTACAGTTACCAATGCTTAACTAGTGAAGCAGCCTATCTCAGCGATCTGCTAT 4450
DB 2755 ttcgttcacatagtttgcgtgactccccctgctgtagagatacactacacatcagggaggt 2814
QY 4451 TCGTTCATGCTATAGTTGCTTGAAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4510
DB 2815 taccatctggccccagtgctgaatgatacgcgcagagaccacgcgtccaggtccagatt 2874
QY 4511 TACCATCTGGCCCCAGTCTGCTGAATGATACCGCGAGACCACCGCTCACCGGCTCCAGATT 4570
DB 2875 tatacgaataaaccagcagccggaagggccgagcgcagaagtggtcctgcacattat 2934
QY 4571 TATCAAGATAAACCAGCCAGCCGAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 4630
DB 2935 ccgcctccatccagctctatattgttcccggaagctagagtaagtagtgcgcagatta 2994
QY 4631 CGGCTCCATCCAGTCTATTAAITGTTCCGCGGAGGCTAGAGTAAGTAGTTCGCCAGTTA 4690
DB 2995 atagttgcgcacactttgtgcatggtctacagcgcagcgtcgtcgtcgtcgtcgtcgt 3054
QY 4691 ATAGTTTGGGCAACGCTTCTTCCTATGCTAGAGCCTGCTGCTGCTGCTGCTGCTGCTG 4750
DB 3055 gtaggttcactccagctccggttcccaacgatcaagcgcaggttacctatgctcccccag 3114
QY 4751 GTATGGCTTCATTCAGCTCCGCTTCCCAACGATCAAGCGAGATTACATGATCCGCCATGT 4810
DB 3115 tctgaaaaaaacggttagtctcctcgtcctccgactcgttgcagagtaagtagtgccg 3174
QY 4811 TGTGAAAAAAGCGGTAGCTCTCTCGTCTCGGTCCTCGATGCTGTTGTCAGAAAGTAAGTTGGCCG 4870
DB 3175 cagttattacactcatgatttgccagcactgcataatctcttactgtcagtcgcatccg 3234
QY 4871 CAGTGTATCACTCATGTTATGGCAGCAGCTGCATAATCTCTTACGTGTCATGCCATCCG 4930
DB 3235 taagatctttctgagctgtagtactcaacacagctatctcgaagatagtagtc 3294
QY 4931 TAAGATGCTTTTCTGTGACTGTGTGAGTACTCAACCAAGTCAATCTGAGATAGTGTATGC 4990
DB 3295 ggcgacagagttgctctggccgcgcgtcaatacagggataataacgcgcacacatagcaga 3354

```

```

QY 4991 GGCACCGAGTGTGCTCTTGGCCGGGTCTAAATAAGGATAAATAAGGATAAATAAGCAGAA 5050
DB 3355 ctttaaaagtgcctcatcattggaaaacgcttcttcggggcgaaaacactctcaagagatctac 3414
QY 5051 CTTTAAAGTGTCTCATCATTTGSAACACGTTCTTTCGGGGGCGAAACCTCTCAAGGATCTTAC 5110
DB 3415 cgcgttgagatccagttcagatgaacccactcgtgcacccactgactcttcagcagctt 3474
QY 5111 CGCTGTTGAGATCCAGTTCGATGTACCCACTCGTGCACCCCACTGATCTTCAGCATCTT 5170
DB 3475 ttacttccaccagcttcttcgggtgagcaaaaaacaggaagcaaaatgcgcgcaaaaag 3534
QY 5171 TTACTTTTCCACAGCGTTTCTGGGTGAGCAAAAAAGAGGAAATGCGCAAAAAAGG 5230
DB 3535 gaataagggcgacacagaaatgttgaatacactcactcactcactcactcactcactcactcact 3594
QY 5231 GAATAAGGGGAGACAGGAAATGTTGAATACTACTACTCTCTCTCTCTCTCTCTCTCTCTCTCT 5290
DB 3595 gcatttatcagggttatgtctcatcaggggacacatatattgaatgatttttadaaaaa 3654
QY 5291 GCATTTATCAGGTTATTGTCTCATGAGCGGATACATATTGAATGATATTAGAAAAATA 5350
DB 3655 acaaaataggggtccgcgcacatttcccccgaagagtgccac 3696
QY 5351 AACAAATAGGGGTTCGCGCACATTTCCCGGAAAAAGTGACAC 5392

```

## RESULT 7

```

ID Q40279 standard: DNA: 4145 BP.
AC Q40279;
DE 02-AUG-1993 (first entry)
DE Sequence of clone pS2gpt-S4.
KW Plasmid; cloning; restriction site; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 1..2226 /*tag= a
FT /*label= pN2gpt-S4
FT /*note= "Posn. 1 corresp. to the first nucleotide
FT G '5'-TGCACACTTT TCGGGGAAAT-3'."
FT misc_feature 2227..2236 /*tag= b
FT /*label= SmaI adaptor
FT 2396..2851 /*tag= c
FT /*label= E. coli gpt gene
FT complement (3081..3323)
FT /*tag= d
FT /*label= vaccinia p7.5 promoter
FT promoter 3358..3451 /*tag= e
FT /*label= S4 of pN2gpt-S4
FT /*note= "oligo P-attP(9)."
FT misc_feature 2237..4145 /*tag= f
FT /*label= pN2gpt-S4

```

```

PN AU9221269-A.
PD 04-MAR-1993.
PF 25-AUG-1992; 021269.
PR 26-AUG-1991; US-750080.
PR 20-JUL-1992; US-914738.
PA (IMMO) IMMUNO AG.
PI Dörner F, Falkner FG, Pfeleiderer M, Scheiflinger F;
DR WPI: 93-126461/16.
PT Modified eukaryotic cytoplasmic DNA virus prodn. * involves
PT direct molecular cloning of modified DNA molecule contg.
PT cytoplasmic DNA virus genome
PS Example: Pages 160-162; 206pp; English.
CC Plasmids pN2gpt-S3A and pN2gpt-S4 comprise expression cassettes
CC with a selective marker. These plasmids were constructed by first
CC making plasmids pN2-gpta and pN2-gptb which contain an E. coli gpt
CC gene driven by the vaccinia virus p7.5 promoter, flanked by
CC several unique restriction sites including NotI. The plasmid

```







QY	33931	GCCTTCATAGCTCAGCTGTAGGTATCTCAGTTCGGGTAGTGCCTGCCTCCAGCT	38990
DB	33888	ggctctatgtcacgaacccccccttcagcccgacccgtcgcccttatccggtaactatcg	3347
QY	33931	GGCGTGTGTGCAGAACTCCCGCGTTCAGCGCGAGCGGTGGCGCTTATCGGTAACTATCG	3950
DB	33488	tcttgaatccaaaccccgtagtaagacagagacttatcgccactggcagcagccactgtagaacag	3407
QY	33951	TCCTTAGTCTCAACCCCGTAAAGACAGACTTATCGGCATCGCAGCAGCCACTGGTAACAG	4010
DB	34088	gattagcagagcagagtatgttaggcggtgcgtacagagtctctgaagtgggtggcctaacta	3467
QY	43211	GATTAGCTAGAGCTGAGGTAATGTAAGCGGCTGCTACAGAGATTCCTTGAAGTGGTGGCTTAACATA	4070
DB	34688	cggctacactaagaagacagagtatttggttatcttgccctctgctgtgaagccgaatttaacctcgg	3527
QY	43711	CGGCTACACTAGAGAGACAGATTTTGTTATCTGGCGCTCTGCTGAAGCAGCTTACCTTCGG	4130
DB	35288	aaaaaagatttggtagctctctgtatccggcgaaacaaaccaccgctggtgaagcgttggtttt	3587
QY	43131	AAAAAGAGTTTGGTAGTCTCTGATCGCGCTAAGCTAAAGCTACCGCTGTACCGGTGGTTTTT	4190
DB	35888	tgtttcacagcagcagattacgcgcagagaaaaagagatctccaagaagatcctcttgatctt	3647
QY	43191	TGTTTGTAACTAGCAGATTACCGCTAGAGAAAGAGATCTCAAGAAAGATCCTTTGATCTT	4250
DB	36488	ttctcagggtctgcagctcaagtgaacgaaacactcaagttcaagggtatttggtcatgag	3707
QY	43511	TCTACCGGCTCTGATGCTCTAGTGAAGCTAATGTAAGAGATTTTGTCATGAG	4310
DB	37088	attatcaaaaaagatcttcaccttagatccttttaaatcaaaaatgaagcttttaaatcaat	3767
QY	43111	ATTATCAAAAGGATCTTCACCTAGATCCTTTTAAATTAAGATGAAGTTTAAATCAAT	4370
DB	37688	ctaaagtatatatgataaacttggtctgcacagttaccaatgcctaatcaatgaggacc	3827
QY	43711	CTAAGTATATATGATTAACCTGGTCTGACAGTATACCAATGCTTAATCACTAGGAGCC	4430
DB	38288	tatctcagcagatctgtctatttggttcacatccatagttgcctgagctgccctcctgtagat	3887
QY	44311	TATCTCAGCCATCTGTCTATTTCCTTCATCTCATAGTTGCTGTGATCCCGCTGTGTAGAT	4490
DB	38888	aaactacatagtcgggaggggtaccatctctggccccagtgctgcaatgatcaccgcgagacc	3947
QY	44311	AACTACGATATAGGAGGAGGTTTATCTCTGGCTGCAAGTGTGCAATGATACCGCGAGAGCC	4550
DB	39488	acgctcaccggtctcagatttatcaacaataaacacagcagccaggaagggccagcagcag	4007
QY	45511	ACGCTGATGCTGCTGATATATATGCAATTAACCTAGCGAGCGCGAGGCGCTGAGCGAG	4610
DB	40088	aaatggctcgcgaactttatccgcctccatccagctctatttaattgttgcggggaagctag	4067
QY	46111	AAGTGGTCTTGCAACTTATCCGCTCCCATCCAGTCTATTAATTGTTCCGCGGAAGCTAG	4670
DB	40688	agtaagtagttgccagtttaagtgttgcgcaacgttgttgcattgctacaggcacgtgt	4127
QY	46711	AGTAAGTAGTTGCGCAGTTAATAGTTTGGCGAACGTTGTGGCATTTGTACAGGCATCGT	4730
DB	41288	gggtacagctcgttcgttttgtagtgccttcattcagctccggttcccaagcatcaaggcg	4187
QY	47311	GGTGTACGCTCGTGTGTTTGGTATGCTTCATTCAGTCCGCTTCCCAACGATCAAGGCG	4790
DB	41888	attacatgatcccccatgttgttaaaaaaacgggttaagctccttgctccgactcgt	4247
QY	47911	AGTTACATGATCCCGCAAGTTGTCAGAAAAGTGCTAGCTCTCTCGGTCTCGGTATCGT	4850
DB	42488	tgctcaagtaagttgaccgcagtgattacactcatgctatggcagcagcactgcataatc	4307
QY	48511	TGTCAGAGTAGTTGGCGCGAGTGTATATCATCTATGTTATGGCAGCACTTGCAATTC	4910
DB	43088	ttctactgataatgcatcgtatgaaga-gctttctgtgactgtagtagtactaaaccaagtc	4367

RESULT 10

RESULTS	ID
	ID Q84694 standard: DNA: 5042 BP.

ID Q84694  
AC Q84694:

DT 29-DEC-1995 (first entry)

DI 29-DEC-1993 (11:53 am) Plasmid glucoamylase Eco RI pGE-

DE plasmid glucoamylase Eco RI pGE.  
KW *Neurospora crassa*, glucoamylase gene,  $\alpha$ -1 promoter.

KW Neurospora  
KW expression

expression c  
synthetic

OS Synthetic.  
PN W09505474-A

PN WO9505474-A.  
PD 23-FEB-1995

PD 23-FEB-1995.  
 DE 15-AUG-1994: 501789

PF 15-AUG-1994: G01789.  
PB 13-AUG-1993: GP-016883

PR 13-AUG-1993; GB-016883.  
PA (UNITE-) UNTV FEEDS

PA (UYLE-) UNIV LEEDS.  
PT Parich in Radford

PI Parish JH, Radford A;  
 PP WPT: 05-000771/12

DR WPI; 95-098771/13.

Regulated glucoamylase promoter

polypeptide(s) in filamentous fungi

PS Claim 11; Fig 6; 36pp; English.

The 5' primer (Q8469U) encompassed the unique Ppum I site at posn.

CC 2163 of the *N. crassa* glucoamylase ORF and the 3' primer (Q84691)

CC contg. an MroI site hybridises at the 3' end of the gla gene. The

5' upstream PCR fragment was amplified and cloned into the Sma I

CC site in a pNEB 193 vector. The clone was named pMro. The remainder

CC of the gla gene was inserted by digestion of the glucamylase clone

CC pGla-Xho I (Q84692). This plasmid contains the entire gla gene

CC however the downstream unsequenced and non-transcribed area was

CC deleted. pGla-Xho I was digested with Sac I and Ppum I and the

CC fragment ligated into the Sac I/PvuII sites of pMro 1. The Sac I

CC site of pGla-Xho I was derived from the linker and not from the

CC coding region of glucoamylase consequently, no glucoamylase

CC sequence was deleted (see Q84593, pGla-Mro I). In an

CC attempt to increase transcription efficiency, 1575 bp were deleted

CC from the glucoamylase ORF, creating the plasmid pGE (plasmid

CC glucoamylase, Eco RI). Deposits of plasmids pGLA-Xho

CC (Q84692), pGLA-Mro I (Q84593), pGE (Q84694) have been made and

U  
U



QY 5271 CTTTTTCAATATTATTGAAGCATTATATCAGGGTATTATGTCTCATGAGCGATACATATT 5330  
Db 4978 tgaattgtattgaataaataaataaggggttccgcacatttccgcaaaagtacc 5037  
QY 5331 TGAATGATTATAGAAATAAACAATAAGGGTTCCTGGGACATTTTCCCGGAAAGTGCC 5390  
Db 5038 ac 5039  
QY 5391 AC 5392

RESULT 11  
ID T43794 standard; DNA: 5356 BP.  
AC T43794:  
DI 18-FEB-1997 (first entry)  
DE plasmid PRIPHAT (rat insulin promoter-human IAPP transgene).  
KW type II diabetes mellitus; transgenic animal model; pancreas;  
KW islet; beta cell; islet amyloid polypeptide; IAPP; hyperglycaemia;  
KW glycosuria; diabetic glomerulosclerosis; plasmid PRIPHAT; ds;  
KW cyclic.  
OS Chimeric Homo sapiens;  
OS Chimeric Rattus sp.;  
OS Chimeric synthetic.  
FH Key Location/Qualifiers  
FI promoter 719..1594  
FI /tag= a  
FI /label= RIP-II-promoter  
FI /note= "rat insulin II promoter"  
FI cds 1589..1866  
FI /tag= b  
FI /label= IAPP  
FI /note= "human IAPP coding sequence (Claim 7)"  
FI intron 1861..2580  
FI /tag= c  
FI /label= Albumin\_intron  
FI /note= "human albumin intron I"  
FI terminator 2575..3119  
FI /tag= d  
FI /label= Poly-A  
FI /note= "human GAPDH gene poly-A and RNA  
FI termination region"  
FI WO9637612-A1.  
PN 28-NOV-1996. IB0371.  
PF 24-APR-1996. IB0371.  
PR 23-MAY-1995. US-446935.  
PA (F12 ) PFIZER INC.  
PI Carty MD, Kreutter DK, Soeller WC;  
PI WFI: 97-021221/02.  
PT Recombinant DNA for expression of islet amyloid polypeptide - to  
PT develop prods. for use in diagnosis, study and treatment of  
PT disorders, e.g. diabetes and obesity  
PS Claim 6; Page 21-23; 46pp; English.  
CC Plasmid PRIPHAT (T43794) allows pancreatic beta cell-specific  
CC expression of human islet amyloid-associated polypeptide (IAPP)  
CC under the regulation of the rat insulin II promoter in both cell  
CC lines and in transgenic animals (models of type II diabetes  
CC mellitus). It was constructed by PCR amplification (see also  
CC T43795-804) of the relevant DNA fragments and their insertion into  
CC a pBluescript SK(-) framework. The transgenic animals develop  
CC amyloid plaque deposits in islet cells, fasting hyperglycaemia,  
CC glycosuria and diabetic glomerulosclerosis. They can be screened  
CC for treatments that modulate disease progression. Cell lines can  
CC be screened for treatments that inhibit human IAPP expression.  
SQ Sequence 5356 BP; 1368 A; 1294 C; 1274 G; 1420 T;  
Query Match: 40.8%; Score 2202; DB 25; Length 5356;  
Best Local Similarity 100.0%; Pred No. 0.00e+00;  
Matches 2202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3155 ccagctttgttcccttttagtgaggttaattgcgcgcttgccgtaataatcggtcatagc 3214  
QY 3191 CCAGCTTTTGTCCCTTTAGTACGGGTTAATTTCGGCGCTTGGCGTAAATCATGTCATAGC 3250

Db 3215 tgtttcctgtgtgaattgtattccgctcacaattccacacacatacagcgccggaaca 3274  
QY 3251 TGTTCCTGTGTGAATTTGTATCCGCTCACAATTCCACACATATACGAGGCGGAAAGCA 3310  
Db 3275 taaagtgtaaagcctgggtggtcetaatagtgagctaaetccacattaatgttgcgtc 3334  
QY 3311 TAAAGTGTAAAGCCTGGGCTGCTTAATGAGTGAAGTAACTACATTAATTGCTTGGCT 3370  
Db 3335 cactgcgcgctttccagtcgggaaacotcgtcccaactccatattaatcggccaac 3394  
QY 3371 CACTGCCGCTTTCAGTCGGGAAACCTGCTGCTAGCTGCTGCTTAATTAATGAGTGAAGCA 3430  
Db 3395 gcgcgggagagcggtttgcgtattggcgctcttccgcttccctcactcactcactcgc 3454  
QY 3431 GCGCGGGGAGAGGCGGTTTGGGTATTGGGCGCTTTCGCTTCCGCTACATGACTCG 3490  
Db 3455 tgcgtcgtgttgcgtcggcgagcggtatcagctcactcactcaaaagcggttaatacgt 3514  
QY 3491 TCGGCTCGGCTTTCGGCTGCGGAGCGGTATCAGTCACTCAAAAGCGGTAATAGGT 3550  
Db 3515 tatccacagatcaggggataaacccaggaagaacatgtgacaaaggcccaaaagg 3574  
QY 3551 TATCCACAGATCAGGGGATTAACCCAGGAAGAACATGTGAGCAAAAGGCCGCAAGAG 3610  
Db 3575 ccaggaaacctaaaggcgcttgcgtgcgttttccataggtccgccccccctgaag 3634  
QY 3611 CCAGGAACCGTAAAAAGCGCGCTTGTGCGCTTTTTCATAGGCTCCGCGCCCTCAG 3670  
Db 3635 agcatcacaataatcgagcgctcaagtcagaggtggcgaaacccgacagactataagat 3694  
QY 3671 AGCATCACAATAATCGAGCGCTCAAGTCAGAGGTGGCGAAACCCGACAGSACTATAAGAT 3730  
Db 3695 accagcgcttcccccctggaaagctccctcgtggcgctcctgttccgacccctgcctta 3754  
QY 3731 ACCAGCGCTTTCCCGCTGGAAAGCTCCCTCGTGGCGCTCTCTCTGTCCGACCTGCCGCTTA 3790  
Db 3755 ccggatacctgcgccttttcccttcgggaagcggtggcgcttctctatagctcacgt 3814  
QY 3791 CCGGATACCTGCGGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCATAGCTCAGCT 3850  
Db 3815 gtggtatctcagtcgggttagtgccttcgctccaaagctgggtgtgtgcagaaacccc 3874  
QY 3851 GTAGGTATCTCAGTTCCGCTGTAGGTGCTGCTCCAAAGCTGGGCTGTGTGCACGAGACCC 3910  
Db 3875 ccgttcagcccgacgcgtgcgccttaaccgtaactcctcttgcacaaaccccgataa 3934  
QY 3911 CGGTTCCAGCCGACCGCTGCGGCTTATCCGTAATCTGCTTGTAGTCCGACCCGCTAA 3970  
Db 3935 gacagacttatcgccactgcagcagccactgtaacaggtatgacagagcgaggtatg 3994  
QY 3971 GACAGACTTATCGCTCAGTGGTACGAGTCACTGTAAACAGSATTAGTAAAGTGAAGTATG 4030  
Db 3995 tagcggtgctcacaggtcttgaagtggtggccctaaactacggtacactacagagacag 4054  
QY 4031 TAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACCTACGCTACACTAGAGAGACAG 4090  
Db 4055 tatttggtatctgcgctcgtcgtgaagccagttaccttcgaaaaagagttggtagctctt 4114  
QY 4091 TATTGTTATCTGCTGCTGCTGAAGCCAGTACCTTCGAAAAAGAGTTGTTAGTCTTT 4150  
Db 4115 gatcgggcaacaaacacacgcgtgtagcggtggtttttttgttgcgaacacacaaatta 4174  
QY 4151 GATCGGGCAACAAACACCGCTGTTAGCGGTGCTTTTGTGTTGAAAGCAGCAGATTA 4210  
Db 4175 ccgcagaaaaaaaggatctcaaaagatccttgcattcttctacgggtctgacccctc 4234  
QY 4211 CCGCGAGAAAAAAGATCTCAAGAAATCCTTTTATCTTTTCTACGCGGTCTGAGCGCTC 4270  
Db 4235 agtgaacgaaaaactcacggttaagggtatttggctcatgattatcaaaaagatcttcca 4294  
QY 4271 AGTGAACGAAAAACTCACGCTTAAGGGATTITGGTCATGATATCAAAAAGAGATCTTCA 4330  
Db 4295 cctaatccttttaataataaaatgaagttttaaatcaatcaatataatataatgaatga 4354















Db	2268	agcttttggctccctttagtgagggttaatttcgaagctttggcgtaatacatggtcctaagctg	2322
Qy	3193	AGCTTTTGTTCCTTTTAGTGAGGGTTAATTTGGCGCTTTGGCTTAATCATGGTCAATAGCTG	3252
Db	2328	ttctctgtgaaattgttatccgtccacaattcacacaacatacagagccggaagcata	2387
Qy	3253	TTCTCTGTGTGAATTTATCGGCTCACAATTCCACACACATACAGNGCGGAAGCAT	3312
Db	2388	aagtgtaaagctgggggtcctaatagttgagctaaactcaataattggttgcgtca	2447
Qy	3313	AASTGTAAAGCCTGGGGTGCCCTAATGAGTGAGCTAACTCAGTAATTTGGTTGGTGCTCA	3372
Db	2448	ctgcgcgctttccagtcggaagaaacctgtcgtaccagctgcatttaataatcgccaacgc	2507
Qy	3373	CTGCCGCTTTCCAGTCTGGGAAACCTGTGCTGCCACCTGCATTAATGAATCGGCCAACGC	3432
Db	2508	gcggggaagcggggtttgcgtattggggaactcttcgcgttctccgtccactgaactgcgtc	2567
Qy	3433	GGGAGTAAGAAGGCTTTGGGTATTTGGCGCTTTGGCGCTTCTCCGCTTCTCGCTCACTGCTCGTG	3492
Db	2568	cgctcagctctcgctgcgcggaagcggtatcagctcaactcaaaaggcggtaatcaggtta	2627
Qy	3493	CGTCCGATAGGTTTGGGTGGGAGTGTATCAGCTCTACTCAAGGCGGTAAATACGGTTA	3552
Db	2628	ttcacagaatcacgggataaacgcagggaagaacatgtgacaaaggccagcaaaagccc	2687
Qy	3553	TCACACAAATCAGGGGATAAAGCGGAGAAACATGTGAGCAAAAGGCCACAAAGGCC	3612
Db	2688	agaaacccgtaaaaagccgcggttgtcggttttccataggctccgcgcccttcagcag	2747
Qy	3613	AGGAACCGTAAAAAGCCGCGTGTGTGGGGTTTTTCCATAGGCTCCGCCGCCCTTGACGAG	3672
Db	2748	cattcacaaaatcgacgcctcaagtcagagtgagcgaaacccgacagactataaagatac	2807
Qy	3673	CATCACAAAATCGACGCTCAAGTCAGAGGTTGGCCAAACCGACAGACACTATAAGATAC	3732
Db	2808	caggcgtttcccctggaaagctccctgtgcgctctcctgttcogaaacctgcgccttacc	2867
Qy	3733	CAGGCTTTTCCTCTGGAAGCTCTCTGTGGCGTCTCTGTTCGACCTGCCCTTACC	3792
Db	2868	ggataccctgtccgcctttccctctcggaagacgtggcgctttctcatagctcacgctgt	2927
Qy	3793	GGATACCTGTCCGCTTTTCTCCCTTCCGGAAGTGTGGCGCTTCTCATAGGCTCAGGCTGT	3852
Db	2928	aggtatctcagttcggtagtgagtcgttcgtccaaagctgggctgtgacagaaaccccc	2987
Qy	3853	AGGTATCTCAGTTCCGTTAGTGCTGTCCCTCCAAAGCTGGGCTGTGTGCAGGAACCCCCC	3912
Db	2988	gttcagcccaacgcgtgcgccttaccgataaactcatcgttttgagtcacaacccgtaaga	3047
Qy	3913	GTTCAGTGGAAGGCTTGGGCTTTATCCGTAATATGCTGTGAGTCCAAACCCGGTAAGA	3972
Db	3048	caagacttatccgacctggcagcagccactggtataacagattagcagagcaggtatgta	3107
Qy	3973	CAGGACTTATCGCCACTGGCAGAGCCACTGCTAACAGGATTAGCAGAGCAGGATGTA	4032
Db	3108	gcgagttctacagagttcttgaagtgttggtcctaactcagctacactagaagacagta	3167
Qy	4033	GGCGGTGCTACAGAGTCTTGAAGTGGTGGCCCTAACTACGCTACACTAGAAGACAGTA	4092
Db	3168	tttgatctcgccctctgctggaagccagttaccttcgaaaaaagttggtagctctga	3227
Qy	4093	TTTGGTATCTCGCTTGTGTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGA	4152
Db	3228	tcggggaacaaacccgcgtggtagcggtggttttttggttgcaagcagcagattacg	3287
Qy	4153	TCCGGCAAAACAACACACCGCTGTGTAGCGGTGTTTTTTGTTTGGCAACACGACATTACG	4212
Db	3288	cgcagaaaaaagagatctcaaaaagatcctttgatcttttctacggggtctgagctcag	3347
Qy	4213	CGCAGAAAAAAGATCTCAAGGAAGATCCTTTGATCTTTTCTTACGGGCTGTGACGCTCAG	4272
Db	3348	tgaacgaacaaactcacgttaagggattttggtcatgaattatcaaaaagatcttcacc	3407

09 5353 AAAATAGGCTTCGGCGGACATTTCGCGAAAAAGTGGCAG 5392

Search completed: Tue Mar 16 03:14:30 1999  
Job time : 1469 secs.

\*\*\*\*\*  
[W][O][R][L][D]  
[D][I][S][T][R][I][B][U][T][I][O][N]  
[R][I][G][H][T][S]  
[B][Y]  
[O][X][F][O][R][D]  
[M][O][L][E][C][U][L][A][R]  
[L][I][M][I][T][E][D]  
[C][O][P][Y][R][I][G][H][T]  
[1][9][9][3]  
\*\*\*\*\*

(TM)

Release 3.1A John F. Collins, BioComputing Research Unit  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution Rights by Oxford Molecular Ltd

MPSrch\_nm n.a. n a database search, using Smith-Waterman algorithm  
Run on: Mon Mar 15 23:40:59 1999; MasPar time 6848.69 Seconds  
Tabular output not generated. 1409.884 Million cell updates/sec

Title: >US-09-020-716-6  
Description: (15392) from US09020716 seq  
Perfect Score: 5392  
N.A. Sequence: 1 CTAATTTGTAAGCGTTAATA...ATTCCCGAAAGTGCCAC 5392  
Comp: GATTACATTCGAATTAT...TAAAGGGCTTTTCACGGTG

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895389244 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: emb1-est56  
Database: genbank-est109  
59gb-est1 6:gb-est10 7:gb-est11 8:gb-est12 9:gb-est13  
10:gb-est14 11:gb-est15 12:gb-est16 13:gb-est17  
14:gb-est18 15:gb-est19 16:gb-est2 17:gb-est20  
18:gb-est21 19:gb-est3 20:gb-est4 21:gb-est5 22:gb-est6  
23:gb-est7 24:gb-est8 25:gb-est9 26:gb-est1 27:gb-est2  
28:gb-est3 29:gb-est4

Statistics: Mean 12.869; Variance 2.781; scale 4.627  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Query Match	Length	ID	Description	Pred. No.	
1	564	12.3	711	26	AG004608 Homo sapiens genomic D	0.00e+00	
2	564	12.3	711	27	AG010947 Homo sapiens genomic D	0.00e+00	
3	557	12.2	721	27	AG011001 Homo sapiens genomic D	0.00e+00	
4	557	12.2	721	25	AG004662 Homo sapiens genomic D	0.00e+00	
5	535	11.8	718	26	AG004363 Homo sapiens genomic D	0.00e+00	
6	535	11.8	718	27	AG010489 Homo sapiens genomic D	0.00e+00	
7	530	11.7	698	26	AG003787 Homo sapiens genomic D	0.00e+00	
8	530	11.7	698	27	AG009976 Homo sapiens genomic D	0.00e+00	
9	524	11.6	696	26	AG009765 Homo sapiens genomic D	0.00e+00	
10	524	11.6	696	26	AG003576 Homo sapiens genomic D	0.00e+00	
11	524	11.6	720	28	AG013858 Homo sapiens genomic D	0.00e+00	
12	524	11.6	720	25	AG005062 Homo sapiens genomic D	0.00e+00	
13	619	11.5	747	26	AG007052 Homo sapiens genomic D	0.00e+00	

C	14	609	11.3	644	28	AQ115013	CIT-HSP-2374C24.TF CIT	0.00e+00
C	15	603	11.2	692	27	AQ040352	CIT-HSP-237K21.TF CIT	0.00e+00
C	16	595	11.2	695	26	AG002183	Homo sapiens genomic D	0.00e+00
C	17	597	11.1	703	26	AG001761	Homo sapiens genomic D	0.00e+00
C	18	598	11.1	748	17	A1124281	1.53 Drosophila 8-12 h	0.00e+00
C	19	593	11.0	640	27	AQ074598	CIT-HSP-2382L21.TF CIT	0.00e+00
C	20	595	11.0	654	26	B78704	CIT78SK-A-880F6.TPC C	0.00e+00
C	21	594	11.0	690	27	AG009464	Homo sapiens genomic D	0.00e+00
C	22	588	10.9	652	28	AQ108633	CIT-HSP-2379M3.TF CIT	0.00e+00
C	23	583	10.8	652	26	C1786	Ciona intestinalis gen	0.00e+00
C	24	583	10.8	666	27	AQ079096	CIT-HSP-2356E24.TF CIT	0.00e+00
C	25	577	10.7	689	28	AG014394	Homo sapiens genomic D	0.00e+00
C	26	566	10.5	642	28	AQ113213	CIT-HSP-2372L22.TF CIT	0.00e+00
C	27	566	10.5	682	27	AQ074693	CIT-HSP-2381L23.TF CIT	0.00e+00
C	28	559	10.4	742	28	AQ108392	CIT-HSP-2383E2.TF CIT	0.00e+00
C	29	557	10.3	615	28	AQ112424	CIT-HSP-2376F2.TF CIT	0.00e+00
C	30	551	10.2	604	26	AG002706	Homo sapiens genomic D	0.00e+00
C	31	550	10.2	634	27	AQ039010	CIT-HSP-2356F2.TF CIT	0.00e+00
C	32	549	10.2	653	27	AQ077504	CIT-HSP-2365K21.TF CIT	0.00e+00
C	33	544	10.1	633	27	AQ040787	CIT-HSP-2336K22.TF CIT	0.00e+00
C	34	545	10.1	661	27	AQ057603	CIT-HSP-2345M2.TF CIT	0.00e+00
C	35	545	10.1	690	27	AQ029499	RPC111-39F1.TP RPC111	0.00e+00
C	36	540	10.0	637	27	AQ077360	CIT-HSP-2365E1.TF CIT	0.00e+00
C	37	534	9.9	594	26	B71605	RPC111-9N1.TP RPC111 H	0.00e+00
C	38	536	9.9	626	28	AQ111342	CIT-HSP-2371M2.TF CIT	0.00e+00
C	39	526	9.8	540	27	AQ074307	CIT-HSP-2382N1.TF CIT	0.00e+00
C	40	531	9.8	549	27	AQ003191	PCF111-19F5.TP PCF111	0.00e+00
C	41	526	9.8	594	26	AG007826	Homo sapiens genomic D	0.00e+00
C	42	527	9.8	629	27	AG011367	Homo sapiens genomic D	0.00e+00
C	43	530	9.8	630	27	AQ003238	RPC111-19F13.TP RPC111	0.00e+00
C	44	530	9.8	743	26	AG002826	Homo sapiens genomic D	0.00e+00
C	45	525	9.7	600	28	AQ111411	CIT-HSP-2373F2.TF CIT	0.00e+00

ALIGNMENTS

RESULT 1  
LOCUS AG004608 711 bp DNA GSS 29-JAN-1998  
DEFINITION Homo sapiens genomic DNA, 21q region, clone: S594BG38. genomic survey sequence.  
ACCESSION AG004608  
NID 92822058  
KEYWORDS GSS.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens DNA, clone: S594BG38.  
REFERENCE 1 (bases 1 to 711)  
AUTHORS Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.  
TITLE Eukaryotae: Metazoa, Chordata, Vertebrata, Mammalia; Eutheria; Primates: Catarrhini, Hominoidea, Homo.  
JOURNAL Published Only in Database (1998) In press  
REFERENCE 2 (bases 1 to 711)  
AUTHORS Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.  
TITLE Direct Submission  
JOURNAL Submitted (29-JAN-1998) to the DDBJ/EMBL/GenBank databases. Masahira Hattori, Kitasato University, Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan (E-mail: hattori@hgc.ims.u-tokyo.ac.jp, Tel: 0427-78-9732, Fax: 0427-78-9561)

FEATURES  
source Location/Qualifiers  
1..711  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/clone="S594BG38"  
/map="21q"  
BASE COUNT 169 a 188 c 172 g 182 t  
ORIGIN  
Query Match 12.3% Score 664; DB 26; Length 711;  
Best Local Similarity 99.6%; Pred. No. 0.00e+00;  
Matches 682; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

```

Db 24 TACATAATACGATACAGGAGAGGTTTACCATCTGGCCCTGAGTGTCTTAATGATATACGGG 87
QY 4486 TACATAATACGATACAGGAGAGGTTTACCATCTGGCCCTGAGTGTCTTAATGATATACGGG 4544
Db 88 AGACGACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 147
QY 4545 AGACGACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 4604
Db 148 GGTGATACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 207
QY 4605 GGTGATACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 4664
Db 208 AGCTAGATACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 267
QY 4665 AGCTAGATACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 4724
Db 268 CATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 327
QY 4725 CATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 4784
Db 328 AAGGATACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 387
QY 4785 AAGGATACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 4844
Db 388 SATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 447
QY 4845 SATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 4904
Db 448 TAATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 507
QY 4905 TAATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 4964
Db 508 TAATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 567
QY 4965 TAATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 5024
Db 568 TAATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 627
QY 5025 TAATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 5084
Db 628 TAATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 686
QY 5085 TAATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 5143
Db 687 GTGACACCAACTGATCTTCAGGATC 711
QY 5144 GTGACACCAACTGATCTTCAGGATC 5168

RESULT 2 AG010947 711 bp DNA GSS 28-JUL-1998
LOCUS Homo sapiens genomic DNA, 21q region, clones: H79G12X91, genomic
DEFINITION survey sequence.
ACCESSION AG010947
KEYWORDS g3347775
SOURCE GSS
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 711)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published only in database (1998) In press
REFERENCE 2 (bases 1 to 711)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (28-Jun-1998) to the DDBJ/EMBL/GenBank databases
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail: hattori@shri.riken.go.jp, Tel: 0427-78-9732,
Fax: 0427-78-9561)
Location/Qualifiers

FEATURES

```

```

source 1..711
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="S594RC38"
/map="21q"
BASE COUNT 159 a 189 c 172 g 182 t
ORIGIN
Query Match 12.4% Score 664 DB 27 Length 711
Best Local Similarity 99.6% Pred. No. 0.00e+00
Matches 682; Conservative 0; Mismatches 0; Indels 4; Gaps 0
Db 28 TAGATAATACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 207
QY 4486 TAGATAATACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 4544
Db 88 AGACGACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 147
QY 4545 AGACGACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 4604
Db 148 GGTGATACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 207
QY 4605 GGTGATACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 4664
Db 208 AGCTAGATACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 267
QY 4665 AGCTAGATACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 4724
Db 268 CATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 327
QY 4725 CATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 4784
Db 328 AAGGATACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 387
QY 4785 AAGGATACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 4844
Db 388 SATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 447
QY 4845 SATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 4904
Db 448 TAATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 507
QY 4905 TAATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 4964
Db 508 TAATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 567
QY 4965 TAATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 5024
Db 568 TAATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 627
QY 5025 TAATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 5084
Db 628 TAATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 686
QY 5085 TAATGCTAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTAC 5143
Db 687 GTGACACCAACTGATCTTCAGGATC 711
QY 5144 GTGACACCAACTGATCTTCAGGATC 5168

RESULT 3 AG011001 721 bp DNA GSS 29-JUL-1998
LOCUS Homo sapiens genomic DNA, 21q region, clones: H79G12X91, genomic
DEFINITION survey sequence.
ACCESSION AG011001
KEYWORDS g3357935
SOURCE GSS
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.

```











Cp	4784	GATCGTTGGGAACGGAGCTGTAATGAAGCTATACCAAGGACGAGGAGGAGGATGACACAGGAT	4726
Db	511	GGCCTGTAGCAATGGCAACAAACGCTGGCAAACTATTAACTGGGCAACTACTTACTACTAG	570
Cp	4725	G-CCTGTAGCAATGGCAACAAACGCTGGCAAACTATTAACTGGGCAACTACTTACTACTAG	4667
Db	571	CTTCCCGGGCAACAAATTAATAGACTGGATGGAGGCGGATAAAGTTGACAGGACCACTCTG	630
Cp	4666	CTTCCCGG-CAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGACAGGACCACTCTG	4608
Db	631	CGCTGGCGCCTTTCGGGCTGGCGTGGTATTATTGCTGATAAATCTGGAGCGGTCAGC-TC	689
Cp	4607	CGCTGGCGCCTT-CCGGTGG-CTGGTATTATTGCTGATAAATCTGGAGCGGTCAGCGTG	4560
Db	690	GGTCTC 695	
Cp	4549	GGTCTC 4544	
RESULT	11		
LOCUS	AG013858	720 bp DNA	GSS 10-SEP-1998
DEFINITION	Homo sapiens genomic DNA, 21q region, clone: 762015SpN14, genomic survey sequence.		
ACCESSION	AG013858		
NID	93560348		
KEYWORDS	GSS.		
SOURCE	Homo sapiens DNA, clone:762015SpN14.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota: Metazoa: Chordata: Vertebrata: Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.		
AUTHORS	1 (bases 1 to 720)		
TITLE	Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.		
JOURNAL	Homo sapiens genomic DNA, chromosome 21q		
REFERENCE	Published Only in DataBase (1998) In press		
AUTHORS	2 (bases 1 to 720)		
TITLE	Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.		
JOURNAL	Direct Submission		
FEATURES	Submitted (10-SEP-1998) to the DBJ/EMBL/GenBank databases.		
Source	Masahira Hattori, Kitasato University, Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan (E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732, Fax:0427-78-9561)		
Location/Qualifiers	1..720		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="21"		
	/clone="762015SpN14"		
	/map="21q"		
BASE COUNT	172 a	178 c	170 g 186 t 14 others
ORIGIN			
Query Match	11.6%; Score 624; DB 28; Length 720;		
Rest Local Similarity	97.2%; Pred No 0.00e+00;		
Matches	671; Conservative	0; Mismatches	13; Indels 6; Gaps 5;
Db	35	AGTCCTCGTGGCGCTCTCGTGTTCGAGCGCTGGCGGATACCTGTCGCGCTT	94
Qy	3751	AGTCCTCGTGGCGCTCTCGTGTTCGAGCGCTGGCGGATACCTGTCGCGCTT	3810
Db	95	CTCCCTTCGGGAACGTTGGCGCTTTCATAGCTACGCTGAGTAICTAGTTCGGTG	164
Qy	3811	CTCCCTTCGGGAACGTTGGCGCTTTCATAGCTACGCTGAGTAICTAGTTCGGTG	3870
Db	155	TAGTTCGTTTCGCTCAAGCTGGGCTGTGTGCAGAACCGCGCTTCAGCGGACCGCTGC	214
Qy	3871	TAGTTCGTTTCGCTCAAGCTGGGCTGTGTGCAGAACCGCGCTTCAGCGGACCGCTGC	3990
Db	215	GCCTATCCGTTAACTATGCTTTAGTCCACCGCGTTAAGACAGCACTTATCGCACTG	274
Qy	3931	GCCTATCCGTTAACTATGCTTTAGTCCACCGCGTTAAGACAGCACTTATCGCACTG	3990



```

/chromosome="21"
/clone="c48B2A23"
/map="21q"

BASE COUNT 194 a 171 c 176 g 194 t 12 others
ORIGIN

Query Match 11.5% Score 619; DB 26; Length 747;
Best Local Similarity 96.3%; Pred. No. 0.00e+00;
Matches 683; Conservative 0; Mismatches 17; Indels 9; Gaps 9;

Db 41 ATCTCGCTCTGCTGAAGCCAGTACCTTCGAAAAAGAGTGGTAGCTCTTGATCGGC 100
QY ATCTCGCTCTGCTGAAGCCAGTACCTTCGAAAAAGAGTGGTAGCTCTTGATCGGC 4158
Db 101 AAACAACACACGCTGTAAGCGTG-ITTTTTTGTTCAGACACAGATACCGGAGA 159
QY AAACAACACACGCTGTAAGCGTG-ITTTTTTGTTCAGACACAGATACCGGAGA 4218
Db 160 AAAAAGGATCTCAAGAAGATCCTTTGATCTTTTACGGGGTCTGACGCTCAGTGAAC 219
QY AAAAAGGATCTCAAGAAGATCCTTTGATCTTTTACGGGGTCTGACGCTCAGTGAAC 4278
Db 220 GAAACTCAGGTTAAGGATTTGGTCATGAGATTATCAAAAGGATCTTACCTAGATC 279
QY GAAACTCAGGTTAAGGATTTGGTCATGAGATTATCAAAAGGATCTTACCTAGATC 4338
Db 280 CTTTAAATTAATAATGAAGTTTAAATCAATCAATCAATATATAGTAAACTTGGTCT 339
QY CTTTAAATTAATAATGAAGTTTAAATCAATCAATCAATATATAGTAAACTTGGTCT 4398
Db 340 GACAGTTACCAATGCTTAATCAGTAGGACCACTATCTCAGCGATCTGTCTATTTCGTCA 399
QY GACAGTTACCAATGCTTAATCAGTAGGACCACTATCTCAGCGATCTGTCTATTTCGTCA 4458
Db 400 TCCATATCTTCTGACTCCCGTGGTGTATATCAATCAATCAATCAATCAATCAATCAAT 458
QY TCCATATCTTCTGACTCCCGTGGTGTATATCAATCAATCAATCAATCAATCAATCAAT 4518
Db 459 GGGCCAGTCTGCAATGATACGGGAGAGGACCGACGCTACCGGCTCCAGATTTATCAGCA 518
QY GGGCCAGTCTGCAATGATACGGGAGAGGACCGACGCTACCGGCTCCAGATTTATCAGCA 4578
Db 519 ATAAACAGGACGCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 577
QY ATAAACAGGACGCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4637
Db 578 CATCCAGCTTATTAAATGTTCCCGGAGAGTACAGTAAAGTAGTTTCCGCGAGTTAATAGT 637
QY CATCCAGCTTATTAAATGTTCCCGGAGAGTACAGTAAAGTAGTTTCCGCGAGTTAATAGT 4695
Db 538 TTGGCAACGTTTTCNCATTTCTGAGGAGGATCTGTTGTTGTTGTTGTTGTTGTTGTTG 597
QY TTGGCAACGTTTTCNCATTTCTGAGGAGGATCTGTTGTTGTTGTTGTTGTTGTTGTTG 4754
Db 598 GGTTCATTTTCAGTTCGGGTTCCCAACCATCAAGGGGTTTAAATGAT 746
QY GGTTCATTTTCAGTTCGGGTTCCCAACCATCAAGGGGTTTAAATGAT 4801

RESULT 14
LOCUS A0115013 644 bp DNA GSS 28-AUG-1998
DEFINITION CIT-HSP-2374C24.TR CIT-HSP Homo sapiens genomic clone 2374C24,
genomic survey sequence.
ACCESSION A0115013
NID g3491134
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 644)
AUTHORS Adams, M. D., Rounsley, S. D., Zhao, S., Bass, S., Linher, K., Golden, K.,
Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, M. and

```

```

Venter, J. C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL Unpublished (1998)
COMMENT
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamas@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/numcen/bac_end_search/bac_end_search.htm.
Seq primer: M13 Reverse
Class: BAC ends.
Location/Qualifiers
1..644
/organism="Homo sapiens"
/clone="2374C24"
/db_xref="taxon:9606"
/sex="Male"
/cell_type="Sperm"
BASE COUNT 135 a 183 c 145 g 179 t 2 others
ORIGIN

Query Match 11.3% Score 609; DB 28; Length 644;
Best Local Similarity 99.8%; Pred. No. 0.00e+00;
Matches 615; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 5 TCACGTGGCGCTGCTTTTACACGCTGCTGACTGGAAAAACCCCTGGCTTACCCACATTAAT 64
Cp 618 TCACGTGGCGCTGCTTTTACACGCTGCTGACTGGAAAAACCCCTGGCTTACCCACATTAAT 559
Db 65 CGCCTTGCAGCACATCCCGCTTTCGCCAGCTGAGTAAATAGCGAGAGGAGGAGGAGGAGGAG 124
Cp 558 CGCCTTGCAGCACATCCCGCTTTCGCCAGCTGAGTAAATAGCGAGAGGAGGAGGAGGAGGAG 459
Db 125 CGCCTTGCAGCACATCCCGCTTTCGCCAGCTGAGTAAATAGCGAGAGGAGGAGGAGGAGGAG 183
Cp 498 CGCCTTGCAGCACATCCCGCTTTCGCCAGCTGAGTAAATAGCGAGAGGAGGAGGAGGAGGAG 439
Db 184 TTAAGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
Cp 438 TTAAGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 379
Db 244 GCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 203
Cp 378 GCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 319
Db 304 CAAGCTCTTAATTCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363
Cp 318 CAAGCTCTTAATTCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 259
Db 364 CCCAAAAAAGCTTGAATAGGCTGATGCTTACAGTGTAGGCGCATCGCCCTGATAGACGCTT 423
Cp 258 CCCAAAAAAGCTTGAATAGGCTGATGCTTACAGTGTAGGCGCATCGCCCTGATAGACGCTT 199
Db 424 TTTCGCGCTTTCAGCTTGGAGTCCACGCTTTCATAGTGGAGCTTTCGCAACAGTGA 483
Cp 198 TTTCGCGCTTTCAGCTTGGAGTCCACGCTTTCATAGTGGAGCTTTCGCAACAGTGA 139
Db 484 ACAACTCAACCTATCTCGCTCTATCTTTTGAATTTAAGGAGTTTTCGCAATTCG 543
Cp 138 ACAACTCAACCTATCTCGCTCTATCTTTTGAATTTAAGGAGTTTTCGCAATTCG 79
Db 544 GCCTATTGTTAAAAAATGAGCTGATTTTAAACAAAATTTAAGCGAATTTTAAACAAATA 603
Cp 78 GCCTATTGTTAAAAAATGAGCTGATTTTAAACAAAATTTAAGCGAATTTTAAACAAATA 19

```





\*\*\*\*\*

W P E R E H (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_nn n.a. - n a database search. using Smith-Waterman algorithm  
Run on: Tue Mar 16 05:59:40 1999; MasPar time 7909 93 Seconds  
Tabular output not generated. 1545.492 Million cell updates/sec

Title: >US-09-020-716-7  
Description: (1-5173) from US00020716 seq  
Perfect Score: 5173  
N A Sequence: 1 CTAATGTGAAGCGTTAATA ATTCCCGGAAAGTGGCAC 5173  
Comp: GATTACATCGCAATTAT TAAAGGGGCTTTCAGCGTG

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 502357 seqs, 1181590623 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: emb157  
1:em\_ba 2:em\_fun 3:em\_htg 4:em\_hum1 5:em\_hum2 6:em\_in  
7:em\_lo 8:em\_lo 9:em\_lo 10:em\_pat 11:em\_ph 12:em\_pl  
13:em\_ro 14:em\_un 15:em\_vi  
Database: genbank110  
16:gb\_bal 17:gb\_bal 18:gb\_htg 19:gb\_in 20:gb\_lo 21:gb\_ov  
22:gb\_pat 23:gb\_ph 24:gb\_pll 25:gb\_pl2 26:gb\_pr1  
27:gb\_pr2 28:gb\_pr3 29:gb\_ro 30:gb\_st 31:gb\_sts 32:gb\_sy  
33:gb\_un 34:gb\_vi

Statistics: Mean 12.510; Variance 7.779; scale 1.621

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Query	Length	ID	Description	Pred. No.
1	2228	43.1	2961	32	ARBL2KSP	pBluescript II KS(-) v	0.00e+00
2	2228	43.1	2961	32	ARBL2KSM	pBluescript II KS(-) v	0.00e+00
3	2228	43.1	2979	32	CVKSLIC	Ligation-independent c	0.00e+00
4	2228	43.1	4289	32	XXU35136	Plasmid pBSL197 cloning	0.00e+00
5	2228	43.1	9655	32	SYNPGR8V	Cloning vector pGR8, c	0.00e+00
6	2228	43.1	9750	32	AF025397	Expression vector pESP	0.00e+00
7	2228	43.1	9760	32	EVU86815	Expression vector pESP	0.00e+00
8	2228	43.1	10765	32	EVU87875	PEST-1 yeast expressio	0.00e+00
9	2224	43.0	2964	32	SYNBLKSPV	BlueScribe KS Plus clo	0.00e+00
10	2224	43.0	2964	32	SYNBLKSMV	BlueScribe KS Minus cl	0.00e+00
11	2224	43.0	2967	32	UC2449	Cloning vector pDIRECT	0.00e+00
12	2224	43.0	4412	32	EVCOPI12N	Expression vector pCOR	0.00e+00
13	2222	43.0	7474	22	A85522	Sequence 26 from Paten	0.00e+00

C	14	2224	43.0	8073	32	AF038556	Cloning vector pFUN, c	0.00e+00
15	2221	42.9	2958	32	ARBLKSM	pBluescript KS(-) vect	0.00e+00	
16	2221	42.9	2958	32	ARBLKSP	pBluescript KS(-) vect	0.00e+00	
17	2213	42.8	4433	32	EVCOPI11N	Expression vector pCOR	0.00e+00	
18	2208	42.7	4267	32	PRS304	Yeast integrative vect	0.00e+00	
19	2208	42.7	4373	32	PRS306	Yeast integrative vect	0.00e+00	
20	2208	42.7	4443	32	PRS303	Yeast integrative vect	0.00e+00	
21	2208	42.7	4783	32	PRS314	Yeast centromere vecto	0.00e+00	
22	2208	42.7	4887	32	PRS316	Yeast centromere vecto	0.00e+00	
23	2208	42.7	4967	32	PRS313	Yeast centromere vecto	0.00e+00	
24	2208	42.7	5056	32	SCU64694	Yeast integrating vect	0.00e+00	
25	2208	42.7	5187	32	U34887	Yeast integrating vect	0.00e+00	
26	2208	42.7	5504	32	PRS305	Yeast integrative vect	0.00e+00	
27	2208	42.7	5693	32	SCU64593	Yeast integrative vect	0.00e+00	
28	2208	42.7	6018	32	PRS315	Yeast centromere vecto	0.00e+00	
29	2202	42.6	2976	32	CVU61229	Cloning vector pKRX, c	0.00e+00	
30	2202	42.6	3426	32	ASPGREEN1	Artificial sequences,	0.00e+00	
31	2205	42.6	3681	22	I15651	Sequence 6 from patent	0.00e+00	
32	2205	42.6	3681	22	A20702	PKSEL5 DNA sequence.	0.00e+00	
33	2202	42.6	3886	32	XXU35139	Plasmid pBSL141 clonin	0.00e+00	
34	2202	42.6	4051	32	XXU35136	Plasmid pBSL121 clonin	0.00e+00	
35	2202	42.6	4285	32	XXU35137	Plasmid pBSL199 clonin	0.00e+00	
36	2202	42.6	4480	32	XXU35134	Plasmid pBSL190 clonin	0.00e+00	
37	2202	42.6	4569	32	XXU35130	Plasmid pBSL142 clonin	0.00e+00	
38	2202	42.6	4755	32	U93713	Cloning vector pFS400	0.00e+00	
39	2202	42.6	4842	32	XXU35125	Plasmid pBSL119 clonin	0.00e+00	
40	2202	42.6	4893	32	U93714	Cloning vector pFS421	0.00e+00	
41	2202	42.6	5035	32	XXU35138	Plasmid pBSL130 clonin	0.00e+00	
42	2205	42.6	5314	22	I15649	Sequence 4 from patent	0.00e+00	
43	2205	42.6	5314	22	A20700	pMTVI DNA sequence.	0.00e+00	
44	2202	42.6	6242	32	U93718	Cloning vector pFS412	0.00e+00	
45	2202	42.6	6868	32	U93719	Cloning vector pFS422	0.00e+00	

#### ALIGNMENTS

RESULT	1	ARBL2KSP	2961 bp	DNA	circular	SYN	10-MAY-1995
LOCUS		pBluescript II KS(-)		vector DNA	phagemid	excised from lambda	
DEFINITION		ZAP11.					
ACCESSION		X52327					
NID		958061					
KEYWORDS		artificial sequence; cloning vector; expression vector; vector.					
SOURCE		cloning vectors.					
ORGANISM		Artificial sequences; Cloning vectors.					
REFERENCE		1 (bases 1 to 2961)					
AUTHORS		Thomas, E.A.					
TITLE		Direct Submission					
JOURNAL		Submitted (20-FEB-1990) to the EMBL/GenBank/DBJ databases. Thomas					
REFERENCE		2 (bases 1 to 2961)					
AUTHORS		Short, J.M., Fernandez, J.M., Sorger, J.A. and Huse, W.D.					
TITLE		Lambda ZAP: a bacteriophage lambda expression vector with in vivo					
JOURNAL		excision properties					
MEDLINE		Nucleic Acids Res. 16 (15), 7583-7600 (1988)					
REFERENCE		3 (bases 1 to 2961)					
AUTHORS		Altburg-Wees, M.A. and Short, J.M.					
TITLE		pBluescript II: gene mapping vectors					
JOURNAL		Nucleic Acids Res. 17 (22), 9494 (1989)					
MEDLINE		90067967					
FEATURES		Location/Qualifiers					
source		1..2961					
misc_feature		/organism="Cloning vector"					
BASE COUNT		747 a 738 c 755 g 721 t					
ORIGIN		1..2961					
		/note="phagemid pBluescriptII KS(-)"					
Query Match		43.18; Score 2228; EB 22; Length 2961;					
Best Local Similarity		100.00; P-Val No 0.00e+00;					

Matches	2228	Conservative	0	Mismatches	0	Indels	0	Gaps	0
Db	734	GTCCAGCTTCAGAGGGGGGGGGTGGTAAATGACGTTTGGTGGGTTTAACTGAGAGGTTAAATGCG	793						
QY	2446	GTGAGCTTCAGAGGGGGGGGGTGGTAAATGACGTTTGGTGGGTTTAACTGAGAGGTTAAATGCG	3005						
Db	744	GGCGTTGGGCTTAATCATGGTTATATCGTTTTCCTGTGTGAAATGTTTATCGGCTCAACAAT	853						
QY	4096	GGCGTTGGGCTTAATCATGGTTATATCGTTTTCCTGTGTGAAATGTTTATCGGCTCAACAAT	4065						
Db	854	TCCACACAAATATACAGAGCGCAAGATATAAAATCTAAAGCTTGGGCTGCTTAATGAGTAG	913						
QY	4066	TCCACACACATACAGAGCGCAAGATATAAAATCTAAAGCTTGGGCTGCTTAATGAGTAG	3125						
Db	914	CTAACTCATATTAATGGGTTTGGTCTACTGCTGCTTTTCAGCTATGAGAAACGCTGTGGTGG	973						
QY	4126	CTAACTCATATTAATGGGTTTGGTCTACTGCTGCTTTTCAGCTATGAGAAACGCTGTGGTGG	3185						
Db	974	CCAGCTGCATTAATCAATGCTCAAAAGCGCGGCGGACAGAGAGGCTTTTGGCTATTGGGGCGCTC	1033						
QY	3186	CCAGCTGCATTAATCAATGCTCAAAAGCGCGGCGGACAGAGAGGCTTTTGGCTATTGGGGCGCTC	3245						
Db	1034	TTCCGCTTTTATGCTATCAATATATCTTGGGCTGCGTGGTCTTGGGCTTGGGAGAGGTAAG	1093						
QY	3246	TTCCGCTTTTATGCTATCAATATATCTTGGGCTGCGTGGTCTTGGGCTTGGGAGAGGTAAG	3405						
Db	1094	AATTCAGCTCAAAAGAGGCTTAATATGCTTATATATCAAGATCAGAGGAGTAACTGAGAGAAAGAA	1153						
QY	3406	AATTCAGCTCAAAAGAGGCTTAATATGCTTATATATCAAGATCAGAGGAGTAACTGAGAGAAAGAA	3465						
Db	1154	CATCTCAATCAAAAGAGCTACAAAAAGGCTAGAGAAAGGTAAGGTAAGGTAAGGTAAGGTAAGG	1213						
QY	3466	CATCTCAATCAAAAGAGCTACAAAAAGGCTAGAGAAAGGTAAGGTAAGGTAAGGTAAGGTAAGG	3425						
Db	1214	TTTTCATAGCTTCGGCGCTCTGACAGAGCATCATCAAAATATGACGCGCTCAAGTCAGAGGTC	1273						
QY	3426	TTTTCATAGCTTCGGCGCTCTGACAGAGCATCATCAAAATATGACGCGCTCAAGTCAGAGGTC	3485						
Db	1274	GTCAAAAGCGTACAGACTATAAAGATATACCAAGGCTTTTCTGCTGTGAAGTCTTCTCTGGTGG	1333						
QY	3486	GTCAAAAGCGTACAGACTATAAAGATATACCAAGGCTTTTCTGCTGTGAAGTCTTCTCTGGTGG	3545						
Db	1334	CTCTGCTTTTCAATCTTGGTGGCTTACAGGATACCTGTGCGGCTTTTCGCTTCGGGAAG	1393						
QY	3546	CTCTGCTTTTCAATCTTGGTGGCTTACAGGATACCTGTGCGGCTTTTCGCTTCGGGAAG	3605						
Db	1394	CGTGGGATTTCTTATAGTATCAAGCTGATAGATATATATGATGATATAGTTCGTCGCTC	1453						
QY	3606	CGTGGGATTTCTTATAGTATCAAGCTGATAGATATATATGATGATATAGTTCGTCGCTC	3665						
Db	1454	CAAGCTGGATCTGTTGTCAGAAAGGCTTTTATAGAGGATACAGGCTTATCGGTTAA	1513						
QY	3666	CAAGCTGGATCTGTTGTCAGAAAGGCTTTTATAGAGGATACAGGCTTATCGGTTAA	3725						
Db	1514	CTATCGCTTGTAGTCCACAGCGGTTAAGACAGATTTATATGCGCATGGTAGAGATATGG	1573						
QY	3726	CTATCGCTTGTAGTCCACAGCGGTTAAGACAGATTTATATGCGCATGGTAGAGATATGG	3785						
Db	1574	TACAGGATTATACAGATGAGTATATAGGCTGGTATACAGATTTCTTGAAGTGGTGGCC	1633						
QY	3786	TACAGGATTATACAGATGAGTATATAGGCTGGTATACAGATTTCTTGAAGTGGTGGCC	3845						
Db	1634	TAACTACGCTACATCAAGAGAGATGATTTGGTATATGCTGCTTGTGAGAGCCAGTTAC	1693						
QY	3846	TAACTACGCTACATCAAGAGAGATGATTTGGTATATGCTGCTTGTGAGAGCCAGTTAC	3905						
Db	1694	CTTGGGAAAAACAGCTTGGTATCTTTTCATCTGGGCAAAACAACGACGCTGTGATGGGTGG	1753						
QY	3906	CTTGGGAAAAACAGCTTGGTATCTTTTCATCTGGGCAAAACAACGACGCTGTGATGGGTGG	4965						
Db	1754	TTTTTTTGTGTCAGAGAGATATATAGGCTGATCAAAAAAGAGATCTCAGAGAGATTCCTTT	1813						
QY	3966	TTTTTTTGTGTCAGAGAGATATATAGGCTGATCAAAAAAGAGATCTCAGAGAGATTCCTTT	4025						





QY 3006 GGGGTTGGGTAATCATGTTGATAGTGTTCCTGTGTGAAATTTGTTATCCGCTCACAA 3065  
Db TCCACACAAATACGAGCGGAGCATAAAGTGTAAAGCCTCGGGTGCCTATAGTCAG 931  
QY TCCACACAAATACGAGCGGAGCATAAAGTGTAAAGCCTCGGGTGCCTATAGTCAG 3125  
Db TAACTCACATTAATTCGGTTGGCTCACTGGCGGCTTCCAGTCGGGAAACCTGCTG 991  
QY TAACTCACATTAATTCGGTTGGCTCACTGGCGGCTTCCAGTCGGGAAACCTGCTG 3185  
Db CGAGCTGCATTAATGATCGGCAAGCGCGGGAGAGCGGTTTCCGTTATGGCGGTC 1051  
QY CGAGCTGCATTAATGATCGGCAAGCGCGGGAGAGCGGTTTCCGTTATGGCGGTC 3245  
Db TTCCGCTTCCTCGCTCACTGACTCGCTGGCTCGGCTTCGGCTGCGGAGCGGTATC 1111  
QY TTCCGCTTCCTCGCTCACTGACTCGCTGGCTCGGCTTCGGCTGCGGAGCGGTATC 3305  
Db AGCTCACTCAAGCGGCTAATACGGTTATCCACAGAAATCAGGGGATACGCGAGGAAGAA 1171  
QY AGCTCACTCAAGCGGCTAATACGGTTATCCACAGAAATCAGGGGATACGCGAGGAAGAA 3365  
Db CATGTGAGCAAAAGGCGCAGCAAAAGCGCAGGAACCGTAAAGCGCGGCTGCTGCGCTT 1231  
QY CATGTGAGCAAAAGGCGCAGCAAAAGCGCAGGAACCGTAAAGCGCGGCTGCTGCGCTT 3425  
Db TTCCATAGGCTCGGCGCCCTCAAGGAGCATCACAAAATCAGGCTCAAGTCAGAGGTG 3485  
QY TTCCATAGGCTCGGCGCCCTCAAGGAGCATCACAAAATCAGGCTCAAGTCAGAGGTG 3485  
Db GCGAAACCCGACAGGACTAATAGATACGAGGTTTCCGCTGGAAGCTCCCTCGTGG 1351  
QY GCGAAACCCGACAGGACTAATAGATACGAGGTTTCCGCTGGAAGCTCCCTCGTGG 3545  
Db CTCTCTGTTCCGAGCGTGGCTTACGAGTACCTTCCGCTTTCCTCTGCGGAG 1411  
QY CTCTCTGTTCCGAGCGTGGCTTACGAGTACCTTCCGCTTTCCTCTGCGGAG 3605  
Db CGTGGCGCTTCTCATAGCTCAGGCTAGGTATCTCAGTTCGGTGTAGTGTGCTGCTC 1471  
QY CGTGGCGCTTCTCATAGCTCAGGCTAGGTATCTCAGTTCGGTGTAGTGTGCTGCTC 3665  
Db CAAGCTGGGCTGTGTGACGAAACCCCGCTTACGCGGAGCGCTGCGCTTATCCGGTAA 1531  
QY CAAGCTGGGCTGTGTGACGAAACCCCGCTTACGCGGAGCGCTGCGCTTATCCGGTAA 3725  
Db CTATCGTCTTTCAGTCCAAACCGGTAAAGACGACTTATCGGCACTGGCAGCGCACTGG 1591  
QY CTATCGTCTTTCAGTCCAAACCGGTAAAGACGACTTATCGGCACTGGCAGCGCACTGG 3785  
Db TAACAGGATTAGCAGAGGAGGTATGAGCGGCTGCTCAGAGTCTTGAAGTGTGCGC 1651  
QY TAACAGGATTAGCAGAGGAGGTATGAGCGGCTGCTCAGAGTCTTGAAGTGTGCGC 3845  
Db TAAGTACGGCTACACTAGAGGAGGTATGAGCGGCTGCTCAGAGTCTTGAAGTGTGCGC 1711  
QY TAAGTACGGCTACACTAGAGGAGGTATGAGCGGCTGCTCAGAGTCTTGAAGTGTGCGC 3905  
Db TTTCGAAAAAGAGTTGGTAGCTCTTATCGGCAACCAACACCGCTGGTAGCGGTGG 1771  
QY TTTCGAAAAAGAGTTGGTAGCTCTTATCGGCAACCAACACCGCTGGTAGCGGTGG 3965  
Db TTTTGTGTTTGAAGCAGGAGTACGCGGCAAAAGAGATCTCAAGAGATCCCTT 1831  
QY TTTTGTGTTTGAAGCAGGAGTACGCGGCAAAAGAGATCTCAAGAGATCCCTT 4025  
Db GATCTTCTCAGGAGTCTCAGGCTCAGTGGAGCAAACTCAGGTTAAGGATTTGGT 1891  
QY GATCTTCTCAGGAGTCTCAGGCTCAGTGGAGCAAACTCAGGTTAAGGATTTGGT 4085  
Db CATGAGATTATCAAAAGGATCTTCACTAGATCCCTTTTAAATTAAGATTTAA 1951  
QY CATGAGATTATCAAAAGGATCTTCACTAGATCCCTTTTAAATTAAGATTTAA 4145

Db ATCAATCTAAAGTATATATGATTAAGTGTGCTGACAGTTACCAATGCTTAAATCASTGA 2011  
QY ATCAATCTAAAGTATATATGATTAAGTGTGCTGACAGTTACCAATGCTTAAATCASTGA 4205  
Db GGACCTATCTCAGGAGTCTGCTATTTTCGTTTCATCCCATAGTGTGCTGACCTCCCGCTG 2071  
QY GGACCTATCTCAGGAGTCTGCTATTTTCGTTTCATCCCATAGTGTGCTGACCTCCCGCTG 4265  
Db GTAGATACTAGATACGAGGAGGCTTACCATCTGCGCCCASTGTGTAATGAIAACGAG 2131  
QY GTAGATACTAGATACGAGGAGGCTTACCATCTGCGCCCASTGTGTAATGAIAACGAG 4335  
Db AGACCCAGCTCAGCGGCTCGAGATTTATCACTAATAAGCAGCAGCGGAAAGGAGGGA 2191  
QY AGACCCAGCTCAGCGGCTCGAGATTTATCACTAATAAGCAGCAGCGGAAAGGAGGGA 4385  
Db AGACCCAGCTCAGCGGCTCGAGATTTATCACTAATAAGCAGCAGCGGAAAGGAGGGA 2251  
QY AGACCCAGCTCAGCGGCTCGAGATTTATCACTAATAAGCAGCAGCGGAAAGGAGGGA 4445  
Db AGCTAGAGTAAAGTAGTTCGCGAGTTAATAGTTTGGCAACGTTTGGCATTTGCTACAGG 2311  
QY AGCTAGAGTAAAGTAGTTCGCGAGTTAATAGTTTGGCAACGTTTGGCATTTGCTACAGG 4505  
Db CATGTGCTGCTCAGCTTCGCTTGGTATGCTTCACTAGCTCCGCTTCCCAACGATC 2371  
QY CATGTGCTGCTCAGCTTCGCTTGGTATGCTTCACTAGCTCCGCTTCCCAACGATC 4585  
Db AAGCGAGTTACATATGCTGCTTGGTATGCTTCACTAGCTCCGCTTCCCAACGATC 2431  
QY AAGCGAGTTACATATGCTGCTTGGTATGCTTCACTAGCTCCGCTTCCCAACGATC 4685  
Db GATCGTTGTGAGAAATAGTGTATGCGGAGCGGAGTTTGGTATGCTTCACTAGCTCCG 2491  
QY GATCGTTGTGAGAAATAGTGTATGCGGAGCGGAGTTTGGTATGCTTCACTAGCTCCG 4685  
Db TAATCTCTTACTGCTATGCTGCTTAAAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCT 2551  
QY TAATCTCTTACTGCTATGCTGCTTAAAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCT 4745  
Db CAAGTCACTTCTGAGAAATAGTGTATGCGGAGCGGAGTTTGGTATGCTTCACTAGCTCC 2611  
QY CAAGTCACTTCTGAGAAATAGTGTATGCGGAGCGGAGTTTGGTATGCTTCACTAGCTCC 4805  
Db GGATAATACCGGCGCAGATAGCAGAACTTTAAAGTGTGCTTCTGCTGCTGCTGCTGCTG 2671  
QY GGATAATACCGGCGCAGATAGCAGAACTTTAAAGTGTGCTTCTGCTGCTGCTGCTGCTG 4855  
Db GGGGCAAACTCTCAAGGATCTTACCGGCTGTTGAGATCCAGTTCGATGTAACCCACTG 2731  
QY GGGGCAAACTCTCAAGGATCTTACCGGCTGTTGAGATCCAGTTCGATGTAACCCACTG 4925  
Db TGCACCACTGATCTTCAAGTCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 2791  
QY TGCACCACTGATCTTCAAGTCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 4985  
Db AGGAAGCAAAATGCGGCAAAAGGAAATAGGCGACACGGAATGTGAATACTCAT 2851  
QY AGGAAGCAAAATGCGGCAAAAGGAAATAGGCGACACGGAATGTGAATACTCAT 5045  
Db ACTCTTCTTCTTCAATATTTGAAGCAATTTATCAGGTTTATGCTGCTGCTGCTGCTGCTG 2911  
QY ACTCTTCTTCTTCAATATTTGAAGCAATTTATCAGGTTTATGCTGCTGCTGCTGCTGCTG 5105  
Db CATATTGATGTTTATGAAAAATTAAGGTTTCCGCGCAGATTTCCCGGAA 2971  
QY CATATTGATGTTTATGAAAAATTAAGGTTTCCGCGCAGATTTCCCGGAA 5165  
Db AGTCCAC 2979  
QY AGTCCAC 5173











```

QY 4866 TAATTCCTTACTGTCATGCAATCCGTAAGATAGCTTTCTGTGACTGGTGTAGTACTCAAC 4745
D 9220 CAAGTCATCTGAGATAGTGTATCGGCGACGAGTGTCTCTTCCCGGGGTCAATACG 9279
QY 4746 CAAGTCATCTGAGATAGTGTATCGGCGACGAGTGTCTCTTCCCGGGGTCAATACG 4805
D 9280 GGATAATACGCGCCACATAGCAGAACTTTAAAGTGTCTATCTTGGAAACGTTCTTC 9339
QY 4806 GGATAATACGCGCCACATAGCAGAACTTTAAAGTGTCTATCTTGGAAACGTTCTTC 4865
D 9340 GGGCGCAAACTCTCAGGAATCTTACGCTGTGTGATGATGATGATGATGATGATGATG 9399
QY 4866 GGGCGCAAACTCTCAGGAATCTTACGCTGTGTGATGATGATGATGATGATGATGATG 4925
D 9400 TGCACCAACTGATCTTACGATCTTTTACTTTTACCGAGCTTTCTGGGTGAGCAAAAC 9459
QY 4926 TGCACCAACTGATCTTACGATCTTTTACTTTTACCGAGCTTTCTGGGTGAGCAAAAC 4985
D 9450 AGGAAGCAAAATGCCGCAAAAGGAATAGGGCGGACACGGAATGTTGAATACTCAT 9519
QY 4986 AGGAAGCAAAATGCCGCAAAAGGAATAGGGCGGACACGGAATGTTGAATACTCAT 5045
D 9520 ACTCTTCTCTTCAATATTAATTAAGCATTTATCAGGTTATGTCATCAGCGGATA 9579
QY 5046 ACTCTTCTCTTCAATATTAATTAAGCATTTATCAGGTTATGTCATCAGCGGATA 5105
D 9580 CATATTTGAATGATTTAGAAAAATAACAAATAGGGGTTCGCGGCACATTTCCCGCGAAA 9639
QY 5106 CATATTTGAATGATTTAGAAAAATAACAAATAGGGGTTCGCGGCACATTTCCCGCGAAA 5165
D 9640 AGTCCAC 9647
QY 5166 AGTCCAC 5173

RESULT 5 AF025397 9750 bp DNA SYN 17-DEC-1997
LOCUS Expression vector pESP-3, complete sequence.
DEFINITION Expression vector pESP-3, complete sequence.
ACCESSION AF025397
NID q2689634
KEYWORDS Expression vector pESP-3.
SOURCE Expression vector pESP-3.
ORGANISM Expression vector pESP-3.
REFERENCE 1 (bases 1 to 9750)
AUTHORS Lu,Q.
TITLE Direct Submission
JOURNAL Submitted (16-SEP-1997) Marketing Analysis, Stratagene, 11011 North
Torrrey Pines Road, La Jolla, CA 92037, USA
COMMENT On Dec 17, 1997 this sequence version replaced gi:2623066.
FEATURES
    source
        location/Qualifiers
            1..9750
                /organism="Expression vector pESP-3"
                /specific_host="Schizosaccharomyces pombe"
                /db_xref="taxon:67775"
                /plasmid="pESP-3"

BASE COUNT 2919 a 1959 c 1983 g 2889 t
ORIGIN
    Query Match 43.1%; Score 2228; DB 32; Length 9750;
    Best Local Similarity 100.0%; Pred. No. 0.00e+00;
    Matches 2228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

D 1273 GTGCACTTTTCGGGAAATGTGGGGAACCCCTATTGTTTATTTTCTAAATACATT 1332
C 5173 GTGCACTTTTCGGGAAATGTGGGGAACCCCTATTGTTTATTTTCTAAATACATT 5114
D 1333 CAATATGATATCCGCTCATGAGCAATTAACCCCTGATAAATGCTTCAATATTAATAA 1392
C 5113 CAATATGATATCCGCTCATGAGCAATTAACCCCTGATAAATGCTTCAATATTAATAA 5054
D 1393 GGAAGAGTATGATTAACAATTCGGTGTGCGCCCTATTTCCTTTTGGCGGATTTT 1452

```

```

Cp 5053 GGAAGAGTATGATTAACAATTCGGTGTGCGCCCTATTTCCTTTTGTGGTATTTT 4964
D 1453 GCTTCTCTCTTTTGTCTACCCAGAAACGCTGGTGAAGTAAAGATGCTCAAGATCAGT 1512
Cp 4993 GCTTCTCTCTTTTGTCTACCCAGAAACGCTGGTGAAGTAAAGATGCTCAAGATCAGT 4934
D 1513 TGGTGCAGAGTGGGTTACATCGAACTGGATCTCAAGCGGCTAAGATCTTCAAGATT 1572
Cp 4933 TGGTGCAGAGTGGGTTACATCGAACTGGATCTCAAGCGGCTAAGATCTTCAAGATT 4874
D 1573 TTGCCCCCAAGAAACGTTTTCATGATGAGCACTTTTAAATTTCTGATGTGGCGGG 1632
Cp 4873 TTGCCCCCAAGAAACGTTTTCATGATGAGCACTTTTAAATTTCTGATGTGGCGGG 4814
D 1633 TATTATCCCGTATTGACGGCGGCAAGCAACTCGTGGCGGATACACTATTCTCAGA 1692
Cp 4813 TATTATCCCGTATTGACGGCGGCAAGCAACTCGTGGCGGATACACTATTCTCAGA 4754
D 1693 ATGACTTGGTTGAGTACTCACCAGTCACAGAAAGCATCTTACGGATGGCATGACAGTAA 1752
Cp 4753 ATGACTTGGTTGAGTACTCACCAGTCACAGAAAGCATCTTACGGATGGCATGACAGTAA 4694
D 1753 GAGAAATATGCAATGCTGTCATTAACCATAGTATTAATTTGCGGCAACTTACTTTGTA 1812
Cp 4693 GAGAAATATGCAATGCTGTCATTAACCATAGTATTAATTTGCGGCAACTTACTTTGTA 4634
D 1813 CAACGATCGAGGACCGCAAGGAGCTTAACCGCTTTTTCACAACTATGGGGATCATGTAA 1872
Cp 4633 CAACGATCGAGGACCGCAAGGAGCTTAACCGCTTTTTCACAACTATGGGGATCATGTAA 4574
D 1873 CTGCGCTTGGTGGTGGAAACCGGAGCTGAATGAAGCATACCAAAACGAGGCTGACA 1932
Cp 4573 CTGCGCTTGGTGGTGGAAACCGGAGCTGAATGAAGCATACCAAAACGAGGCTGACA 4514
D 1933 CCAGATGCTGTAGCAATGCGCAAAACGCTTTCGCAAACTATTAATTTGCGGCAACTT 1992
Cp 4513 CCAGATGCTGTAGCAATGCGCAAAACGCTTTCGCAAACTATTAATTTGCGGCAACTT 4454
D 1993 CTCTAGCTTCCCGGCAACAAATTAATAGACTGGATGAGCGGATAAAGTTGAGGAGCCAC 2052
Cp 4453 CTCTAGCTTCCCGGCAACAAATTAATAGACTGGATGAGCGGATAAAGTTGAGGAGCCAC 4394
D 2053 TTCTGGCTGCGGCTTCCCGGCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2112
Cp 4393 TTCTGGCTGCGGCTTCCCGGCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4334
D 2113 GTGGGTCTGGGCTATCATTTGAGGACCTGCGGCGAGATGCTAAGGCTTCCGCTATCGTAG 2172
Cp 4333 GTGGGTCTGGGCTATCATTTGAGGACCTGCGGCGAGATGCTAAGGCTTCCGCTATCGTAG 4274
D 2173 TTATCTACAGTACCGGAGTCAAGCAACTATGATGATGATGATGATGATGATGATGATG 2232
Cp 4273 TTATCTACAGTACCGGAGTCAAGCAACTATGATGATGATGATGATGATGATGATGATG 4214
D 2233 TAGTGCTCTACTGATTAAGCATGCTGATGATGATGATGATGATGATGATGATGATGATG 2292
Cp 4213 TAGTGCTCTACTGATTAAGCATGCTGATGATGATGATGATGATGATGATGATGATGATG 4154
D 2293 AGATTGATTTAAACTTCATTTTAAATTTAAAGGATGATGATGATGATGATGATGATGATG 2352
Cp 4153 AGATTGATTTAAACTTCATTTTAAATTTAAAGGATGATGATGATGATGATGATGATGATG 4094
D 2353 ATCTCATGACCAAAATCCCTTAACGTTAGTTTTCGTTCCACTGAGCGTCAGACCCGCTAG 2412
Cp 4093 ATCTCATGACCAAAATCCCTTAACGTTAGTTTTCGTTCCACTGAGCGTCAGACCCGCTAG 4034
D 2413 AAAAGATCAAGAGATCTTCTTGAGATCTTTTTCGCGGTAATCTGCTGCTGCTGCTGCTG 2472
Cp 4033 AAAAGATCAAGAGATCTTCTTGAGATCTTTTTCGCGGTAATCTGCTGCTGCTGCTGCTG 3974
D 2473 CAAAAAACCCACCCCTACGAGCGGTGTTGTTTCCCGGATCAAGAGCTACCAACTCTT 2532

```



Cp 4993 GCTTCTCTGTTTTCCTACCCAGCAAAACCTGGTGAAGTAAAGATGCTGAAGATCAGT 4934  
Db 1513 TGGGTACAGAGTGGGTTTACATCGAAGTGGATCTCAACAGCGGTAAAGATCCTTCAGAGTT 1572  
Cp 4933 TGGGTACAGAGTGGGTTTACATCGAAGTGGATCTCAACAGCGGTAAAGATCCTTCAGAGTT 4874  
Db 1573 TTGGCCCGCAAGAACGTTTTCGAATGATGAGCACTTTTAAAGTTCTGTGTGTGGCGGG 1632  
Cp 4873 TTGGCCCGCAAGAACGTTTTCGAATGATGAGCACTTTTAAAGTTCTGTGTGTGGCGGG 4814  
Db 1633 TATTATCCCGTATTGACGCGCGGCAAGAGCAACTCGTGGCGGATACACTATTCTCAGA 1692  
Cp 4813 TATTATCCCGTATTGACGCGCGGCAAGAGCAACTCGTGGCGGATACACTATTCTCAGA 4754  
Db 1693 ATGACTTGGTTGAGTACTCACCAGTACAGAAAGCACTCTACGGATGGCATGACAGTAA 1752  
Cp 4753 ATGACTTGGTTGAGTACTCACCAGTACAGAAAGCACTCTACGGATGGCATGACAGTAA 4694  
Db 1753 GAGAATTTGAGTCTGCGCATTAACATGAGTGAACACTTGGGCCAATCTTCTGA 1812  
Cp 4693 GAGAATTTGAGTCTGCGCATTAACATGAGTGAACACTTGGGCCAATCTTCTGA 4634  
Db 1813 CAACGATCGGAGACCGAAGGAGCTAACCGCTTTTTCACAACATGSGGATCATGTAA 1872  
Cp 4633 CAACGATCGGAGACCGAAGGAGCTAACCGCTTTTTCACAACATGSGGATCATGTAA 4574  
Db 1873 CTCGCCCTTATGCTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGGTGACA 1932  
Cp 4573 CTCGCCCTTATGCTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGGTGACA 4514  
Db 1933 CCAGATGCTGTAGCAATGGCAACAGCTTCGCAAACTATTAACTGGCAACTACTTA 1992  
Cp 4513 CCAGATGCTGTAGCAATGGCAACAGCTTCGCAAACTATTAACTGGCAACTACTTA 4454  
Db 1993 CTCTAGCTTCCCGGCAACAAATTAATAGACTGGATGAGCGCGATAAAGTTCCAGACCCAC 2052  
Cp 4453 CTCTAGCTTCCCGGCAACAAATTAATAGACTGGATGAGCGCGATAAAGTTCCAGACCCAC 4394  
Db 2053 TTCTGCGCTCGGCCCTTCGCGCTGGCTGTTTATTGCTGATAAATCTGGAGCCGCTGAGC 2112  
Cp 4393 TTCTGCGCTCGGCCCTTCGCGCTGGCTGTTTATTGCTGATAAATCTGGAGCCGCTGAGC 4334  
Db 2113 GTGGGTCTCGCGTATCATTTGAGCACTGGGCGGAGATGGTAAGCCCTCCGCTATCGTAG 2172  
Cp 4333 GTGGGTCTCGCGTATCATTTGAGCACTGGGCGGAGATGGTAAGCCCTCCGCTATCGTAG 4274  
Db 2173 TTATCTACAGCGCGGAGTCAAGCAACTATGGATGAACGAATAGACAGATCGCTGAGA 2232  
Cp 4273 TTATCTACAGCGCGGAGTCAAGCAACTATGGATGAACGAATAGACAGATCGCTGAGA 4214  
Db 2233 TAGTGCTCTCACTGATTAAAGCAATTTGTAAGTCTGACAGCAAGTTTACTCATATACTTT 2292  
Cp 4213 TAGTGCTCTCACTGATTAAAGCAATTTGTAAGTCTGACAGCAAGTTTACTCATATACTTT 4154  
Db 2293 AGATTGATTAAAGCAATTTTAAATTTAAAGGATCTAGTGGAAGATCCTTTTGTATA 2352  
Cp 4153 AGATTGATTAAAGCAATTTTAAATTTAAAGGATCTAGTGGAAGATCCTTTTGTATA 4094  
Db 2353 ATCTCATGACCAAAATCCCTTTAAGCTGAGTTTTCGTTCCACTGAGCGCTCAGACCCCGTAG 2412  
Cp 4093 ATCTCATGACCAAAATCCCTTTAAGCTGAGTTTTCGTTCCACTGAGCGCTCAGACCCCGTAG 4034  
Db 2413 AAAAGATCAAGAGATCTTCTTGAGATCCCTTTTTCGCGCGTAAATCTGCTGCTTGCAAA 2472  
Cp 4033 AAAAGATCAAGAGATCTTCTTGAGATCCCTTTTTCGCGCGTAAATCTGCTGCTTGCAAA 3974  
Db 2473 CAATAAACCCAGCTACCGAGGTTGGTTTCTTCCGCGATCAAGAGTACCACTCTTT 2532  
Cp 3973 CAATAAACCCAGCTACCGAGGTTGGTTTCTTCCGCGATCAAGAGTACCACTCTTT 3914  
Db 2533 TTCCGAAGTAACTGGCTTACGACAGCGCAGATACCAATCTGCTCTTCTAGTGAGC 2592  
Cp 3913 TTCCGAAGTAACTGGCTTACGACAGCGCAGATACCAATCTGCTCTTCTAGTGAGC 3894

Db 2593 CGTAGTTAGGCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGTATA 2652  
Cp 3853 CGTAGTTAGGCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGTATA 3794  
Db 2653 TCCTGTTACGAGTGGCTCTGCCAGTGGCGATAAGTGTGTCTTACCGGTTTGTACTCAA 2712  
Cp 3793 TCCTGTTACGAGTGGCTCTGCCAGTGGCGATAAGTGTGTCTTACCGGTTTGTACTCAA 3734  
Db 2713 GACGATAGTTACCGATTAAGCGCGAGCGCTGGGCTGAAAGGCGGCTCTCTGTACNACG 2772  
Cp 3733 GACGATAGTTACCGATTAAGGCGAGCGGTGGGCTGAAAGGCGGCTCTCTGTACNACG 3674  
Db 2773 CCAGCTTGGAGCGAAGCACTACACCGAACTGAGATACCTACAGCGTGAATGAGAAA 2832  
Cp 3673 CCAGCTTGGAGCGAAGCACTACACCGAACTGAGATACCTACAGCGTGAATGAGAAA 3614  
Db 2833 GCGCCACGCTTCCGAAAGGAGAAAGCGACAGTATCCGCTAAGGCGGAGGCTTCCGAA 2892  
Cp 3613 GCGCCACGCTTCCGAAAGGAGAAAGCGACAGTATCCGCTAAGGCGGAGGCTTCCGAA 3554  
Db 2893 CAGGAGCGCAGGAGGAGCTTCCAGGGGAAACGGCTGGTATCTTTATAGCTCTGCG 2952  
Cp 3553 CAGGAGCGCAGGAGGAGCTTCCAGGGGAAACGGCTGGTATCTTTATAGCTCTGCG 3494  
Db 2953 GGTTCGCCACCTCTGACTTGGCGTTCGATTTTGTATCTGCTAAGGCGGCTTCCGAA 3012  
Cp 3493 GGTTCGCCACCTCTGACTTGGCGTTCGATTTTGTATCTGCTAAGGCGGCTTCCGAA 3434  
Db 3013 TATGAAACCCAGCAAGCAAGCGGCTTTTACGCTTCTGCGCTTTTGTGCTGCTTTG 3072  
Cp 3433 TATGAAACCCAGCAAGCAAGCGGCTTTTACGCTTCTGCGCTTTTGTGCTGCTTTG 3374  
Db 3073 CTCATGTTCTTCTCGGTTATCCCTGATCTGTGATAACCGTATTACCGCTTTG 3132  
Cp 3373 CTCATGTTCTTCTCGGTTATCCCTGATCTGTGATAACCGTATTACCGCTTTG 3314  
Db 3133 AGTAGCTGATACCGCTTCGCGCAGCGCAAGCGCAGCGAGTCACTAGTGAAGG 3192  
Cp 3313 AGTAGCTGATACCGCTTCGCGCAGCGCAAGCGCAGCGAGTCACTAGTGAAGG 3254  
Db 3193 AAGCGAAGAGCGCCCAATACGCAAAACCGCTCTCCGCGGCTTGGCGGCTTCAATTA 3252  
Cp 3253 AAGCGAAGAGCGCCCAATACGCAAAACCGCTCTCCGCGGCTTGGCGGCTTCAATTA 3194  
Db 3253 GCAGCTGGCAGCACAGTTCGCCGACTGAAAGCGGCGAGTGAAGCAATTAATG 3312  
Cp 3193 GCAGCTGGCAGCACAGTTCGCCGACTGAAAGCGGCGAGTGAAGCAATTAATG 3314  
Db 3313 TGAGTTAGCTCACTATTAGGACCCCGCTTTTACACTTTTATGCTTCCGCGCTGTAIG 3372  
Cp 3133 TGAGTTAGCTCACTATTAGGACCCCGCTTTTACACTTTTATGCTTCCGCGCTGTAIG 3074  
Db 3373 TGTGTGGAATGTGAGCGGATAACAATTTACACAGGAAACAGATGACCATGATACG 3432  
Cp 3073 TGTGTGGAATGTGAGCGGATAACAATTTACACAGGAAACAGATGACCATGATACG 3014  
Db 3433 CCAAGCGCAATTAACCTCTACTAAAGGAAACAAAGCTGGGTACCGGCTTCCGCTT 3492  
Cp 3013 CCAAGCGCAATTAACCTCTACTAAAGGAAACAAAGCTGGGTACCGGCTTCCGCTT 2954  
Db 3493 AGGTGAC 3500  
Cp 2953 AGGTGAC 2946

RESULT 8  
LOCUS EV067875 10765 bp DNA circular SYN 16-DEC-1997  
DEFINITION pESP-I yeast expression vector, complete sequence.  
ACCESSION U67875  
NID 92689262  
KEYWORDS  
SOURCE Expression vector pESP-I.

ORGANISM Expression vector pESP-1  
 REFERENCE artificial sequence; expression vectors.  
 AUTHORS 1 (bases 1 to 10765)  
 TITLE Using Schizosaccharomyces pombe as a host for expression and  
 purification of eukaryotic proteins  
 JOURNAL Gene 200 (1-2): 135-144 (1997)  
 MEDLINE 98038984  
 REFERENCE 2 (bases 1 to 10765)  
 AUTHORS Lu Q.  
 TITLE Direct Submission  
 JOURNAL Submitted (22-AUG-1996) R&D, Stratagene, 11011 North Torrey Pines  
 Road, La Jolla, CA 92037, USA  
 REFERENCE 3 (bases 1 to 10765)  
 AUTHORS Grafsky A.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-DEC-1997) R&D, Stratagene, 11011 North Torrey Pines  
 Road, La Jolla, CA 92037, USA  
 REMARK Sequence update by submitter  
 COMMENT pESP-1 yeast expression vector is derived from pBluescript from  
 Stratagene for expression and purification of heterologous genes in  
 Schizosaccharomyces pombe.  
 FEATURES  
 source  
 1..10765  
 /organism="Expression vector pESP-1"  
 /db\_xref="taxon:51339"  
 1401..2261  
 /codon\_start=1  
 /product="beta-lactamase"  
 /db\_xref="PID:q1527193"  
 /translation="MSIQHFRVLVPPFAAFCLPVFAHPETLVKVKDAEDOLGARVGY  
 IELDSGKILLESFRRPEERPMPTSTFKLLCGAVLSRI DAGOEQLGRIRHYSONDLVE  
 YSPVTEKHLDTGMTVEELCSAAITMSDNTAANLLLTIGGPKELTFLHNGDVRTRL  
 DRWPELNEAI PDERDITMPVAMATTLRLKLTGELLILASRQQLIDHWKADKVGAPL  
 LRSAIPAGFVIAKPSGERSGRIIAALGPDKSPRIWVITTSQATMDERNQVIA  
 EIGASLIKHW"  
 5464..5538  
 /codon\_start=1  
 /product="3-isopropylmalate dehydrogenase"  
 /db\_xref="PID:q2689263"  
 /translation="MSAPKKVVI PGPHVCGEITAEATKVLKAI SDVRSNVKFPENII  
 LIIGAAIDATGVLPDEALEASKKVDVLLGAVGPKWGTVGSRPEQSKLLKIRLEQL  
 YANLRPCNFASDLSLDSPIKFOFAKGVDTFVVRELGVGFKRKGDDGVDWDS  
 QYVPEVORITRWAAFMALOHLEPPPIWSLDKANVLASRLWRKTVEETIKNEPPTLK  
 VQHLIDSAAMILIKNPHTHGIIITSNMFGDIIIDEASVIPSGLGLPSASLASLPD  
 KNTAFGLIPECHSGAPDLPKKNVPIATILISAAMMLKLSLNPPEBGKAIEDAVKKVLD  
 AGIPTDLGSGNSTTEVGDAVEEVKKILA"  
 8168..8884  
 /codon\_start=1  
 /product="glutathione S-transferase"  
 /db\_xref="PID:q1527195"  
 /translation="MSPILGYKKIKGLVPTIKLLILEEYEEHLYERDEGDKWRNK  
 KFELGLEFPNLPYV IDGVKLTQSMALIRYIAKKNMLGGCPKRAEISMLEGAVLDI  
 RYGVSRITAYSKDETELKVDPLSKLPENLKMFEEDRLCHTYLNGDVRHPTDPMVLDALD  
 VVLYMDPKLIDAPFKVLCVPRKKRFEATPQIKYFKRSKYIAMPDQWQATPFQSGSDHPKK  
 SDLVPRGSDYKDDDDKSGSG"

BASE COUNT 3276 a 2125 c 2201 g 3163 t

Query Match 43.1% Score 2228; DB 32; Length 10765.  
 Best local similarity 100.0% Pred No 0.00e+00;  
 Matches 2228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1273 GTGGCACTTTTGGGGAATGAGGAGGCAACGGCTATTGTTTATTTTCTAAATACATT 1332  
 |||||  
 Cp 5173 GTGGCACTTTTGGGGAATGAGGAGGCAACGGCTATTGTTTATTTTCTAAATACATT 5114  
 |||||  
 Db 1343 CAATATGATCGCTATAGACAAATAAGCGGTGATAAATGCTTCATATATTGAAAA 1392  
 |||||  
 Cp 5113 CAATATGATCGCTATAGACAAATAAGCGGTGATAAATGCTTCATATATTGAAAA 5054  
 |||||  
 Db 1393 GGAAAGATGAGTATTCAATTTGCTGTGCGCTTATTCGCTTTTGGGGCATTTT 1452  
 |||||

Cp 5053 GGAAGAGTATGAGTATTCAACATTTCGGTGTGAGCTATATACCTTTTCTGCAATTT 4994  
 |||||  
 Db 1453 GCCTTCCTGTTTTCGTCACCCAGAAACCTGTGTCAAACTAAAGATGTGAACATCACT 1512  
 |||||  
 Cp 4993 GCCTTCCTGTTTTCGTCACCCAGAAACCTGTGTCAAACTAAAGATGTGAACATCACT 4944  
 |||||  
 Db 1513 TGGGTGCACGAGTGGGTATACATGCAACTGATGCAACAGCGCTAAGATCTTCAAGCT 1572  
 |||||  
 Cp 4933 TGGGTGCACGAGTGGGTATACATGCAACTGATGCAACAGCGCTAAGATCTTCAAGCT 4874  
 |||||  
 Db 1573 TTGCCCCGGAAGAAGCTTTTCCAAATCATGAGCACTTTTAAAGTCTCTGCTATCTGCGAGG 1632  
 |||||  
 Cp 4873 TTGCCCCGGAAGAAGCTTTTCCAAATCATGAGCACTTTTAAAGTCTCTGCTATCTGCGAGG 4814  
 |||||  
 Db 1633 TATTATCCCGTATGAGAGAGGGAAGAGAAATGCGTGGGCAATATATATTCTTACA 1692  
 |||||  
 Cp 4813 TATTATCCCGTATGAGAGAGGGAAGAGAAATGCGTGGGCAATATATATTCTTACA 4754  
 |||||  
 Db 1693 ATGACTTGGTTGACTACTACCACTCACACAAAGCATCTTACGGATGATCATCATATAA 1752  
 |||||  
 Cp 4753 ATGACTTGGTTGACTACTACCACTCACACAAAGCATCTTACGGATGATCATCATATAA 4694  
 |||||  
 Db 1753 GAGAAATATGCTAGCTGTGCTGATTAACCATGACTGATATAAATCTCGCGGCAACTTATTCTCA 1812  
 |||||  
 Cp 4693 GAGAAATATGCTAGCTGTGCTGATTAACCATGACTGATATAAATCTCGCGGCAACTTATTCTCA 4644  
 |||||  
 Db 1813 CAACGATCGAGAGACCAAGAGAGATTAAGCTTTTCTGCAACATGCGGGATATATATA 1872  
 |||||  
 Cp 4633 CAACGATCGAGAGACCAAGAGAGATTAAGCTTTTCTGCAACATGCGGGATATATATA 4574  
 |||||  
 Db 1873 CTCGCCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACAGACAGCTGACAA 1932  
 |||||  
 Cp 4573 CTCGCCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACAGACAGCTGACAA 4534  
 |||||  
 Db 1933 CCACGATGCTGTAGCAATGAGTAAACAGCTTCGGTAAACTATTAACTGGCCCAACTACTTA 1992  
 |||||  
 Cp 4513 CCACGATGCTGTAGCAATGAGTAAACAGCTTCGGTAAACTATTAACTGGCCCAACTACTTA 4454  
 |||||  
 Db 1993 CTCCTAGCTTCCCGCGCAACAATTAATAGCTGGATGAGCGGATGAAGCTTGAGCAATCAC 2052  
 |||||  
 Cp 4453 CTCCTAGCTTCCCGCGCAACAATTAATAGCTGGATGAGCGGATGAAGCTTGAGCAATCAC 4394  
 |||||  
 Db 2053 TTTCTGCGCTCGGGCTTCGCGGTGTGCTTTATTTATTTCTATATAATCTGAGTGTGATG 2112  
 |||||  
 Cp 4393 TTTCTGCGCTCGGGCTTCGCGGTGTGCTTTATTTATTTCTATATAATCTGAGTGTGATG 4344  
 |||||  
 Db 2113 GTGGGTCTCGCGGTATCTATTGAGGACTGGGGGACAGATGGTAAAGGCTTCCCGTATGCTAG 2172  
 |||||  
 Cp 4333 GTGGGTCTCGCGGTATCTATTGAGGACTGGGGGACAGATGGTAAAGGCTTCCCGTATGCTAG 4274  
 |||||  
 Db 2173 TTATCTACACGCGGGGAGTCAAGCAACTATGGATGAAAGAAATAGACATGCTGATCA 2232  
 |||||  
 Cp 4273 TTATCTACACGCGGGGAGTCAAGCAACTATGGATGAAAGAAATAGACATGCTGATCA 4214  
 |||||  
 Db 2233 TAGTGCGCTCACTATTAAGCATTCGGTAACTGTCAGACCAAGTTTACTCATATATATTCT 2292  
 |||||  
 Cp 4213 TAGTGCGCTCACTATTAAGCATTCGGTAACTGTCAGACCAAGTTTACTCATATATATTCT 4154  
 |||||  
 Db 2293 AGATTGATTTAAACATTCATTTTAAATTTAAAGGATCTAGGTGAAGATCTTTTCTGATA 2352  
 |||||  
 Cp 4153 AGATTGATTTAAACATTCATTTTAAATTTAAAGGATCTAGGTGAAGATCTTTTCTGATA 4094  
 |||||  
 Db 2353 ATCTCATGACCAAAATTCCTTAAAGTGAATTTTGGTTCACATGAGAGTCAAGAGGAG 2412  
 |||||  
 Cp 4093 ATCTCATGACCAAAATTCCTTAAAGTGAATTTTGGTTCACATGAGAGTCAAGAGGAG 4044  
 |||||  
 Db 2413 AAAAGATCAAGGATCTCTCTGAGATCTCTTTTCTGAGGCTAATCTGCTGCTTGCAAA 2472  
 |||||  
 Cp 4033 AAAAGATCAAGGATCTCTCTGAGATCTCTTTTCTGAGGCTAATCTGCTGCTTGCAAA 4974  
 |||||  
 Db 2473 CAAAAAAGCAACCGCTACAGGCTGCTGCTTTGCTGCGATCAAGAGTATCAATCTCTT 2532  
 |||||

Cp 3973 CAAAAAACACCGCTACACGAGGCTGTTGTTGCCGATCAAGAGCTACCAACTCTTT 3914  
 Db 2533 TTCGAGAGTAACTGGCTTCAGCAGAGCCAGATACCAATACTCTCTTCTAGTGTAGC 2592  
 Cp 3913 TTCGAGAGTAACTGGCTTCAGCAGAGCCAGATACCAATACTCTCTTCTAGTGTAGC 3854  
 Db 2593 CGTAGTAGCCACCACTTCAAGAACTCTGTAGCAGCCCTACATACCTCTCTGTCTAA 2652  
 Cp 3853 CGTAGTAGCCACCACTTCAAGAACTCTGTAGCAGCCCTACATACCTCTCTGTCTAA 3794  
 Db 2653 TCTGTATACAGTGGCTGTCTGACAGTGGGATAGTCTGTCTTACCGGTTGACACTAA 2712  
 Cp 3793 TCTGTATACAGTGGCTGTCTGACAGTGGGATAGTCTGTCTTACCGGTTGACACTAA 3734  
 Db 2713 GAGATAGTTACCGGATAGAGGCGAGCGGTGCGGCTGAACGGGGGTTCTGTGCACACAG 2772  
 Cp 3733 GAGATAGTTACCGGATAGAGGCGAGCGGTGCGGCTGAACGGGGGTTCTGTGCACACAG 3674  
 Db 2773 CGAGTTGGAGGAGGAGCTTACCGGAACTGAGTACGCTACAGCGTGAAGTATGAGAAA 2832  
 Cp 3673 CGAGTTGGAGGAGGAGCTTACCGGAACTGAGTACGCTACAGCGTGAAGTATGAGAAA 3614  
 Db 2833 GCGGACGCTTCCGAGGAGGAGGAGGAGGAGGATCCGCTAGCGGCGAGGCTCGGAA 2892  
 Cp 3613 GCGGACGCTTCCGAGGAGGAGGAGGAGGAGGATCCGCTAGCGGCGAGGCTCGGAA 3554  
 Db 2893 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2952  
 Cp 3553 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3494  
 Db 2953 GCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3012  
 Cp 3493 GCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3434  
 Db 3013 TATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3072  
 Cp 3433 TATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3374  
 Db 3073 CTCACATGTTCTTCTGCTGCTTATCCGCTGATCTCTGATTAACCGTATTACCGGCTTG 3132  
 Cp 3373 CTCACATGTTCTTCTGCTGCTTATCCGCTGATCTCTGATTAACCGTATTACCGGCTTG 3314  
 Db 3133 AGTGAGTGTATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3192  
 Cp 3313 AGTGAGTGTATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3254  
 Db 3193 AAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3252  
 Cp 3253 AAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3194  
 Db 3253 GCAGCTGGCAGCAGAGGTTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3312  
 Cp 3193 GCAGCTGGCAGCAGAGGTTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3334  
 Db 3313 TGAGTTAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 3372  
 Cp 3333 TGAGTTAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 3074  
 Db 3373 TGTGTGAAATTTGAGCGGATACAAATTTACACAGGAAACAGCTATGACCAATGATTACG 3432  
 Cp 3073 TGTGTGAAATTTGAGCGGATACAAATTTACACAGGAAACAGCTATGACCAATGATTACG 3014  
 Db 3433 CCAAGCGGCAATTAAGCTTCTACTAAGGAGCAACAAAGCTGGTACCGGGGCCCCCTCG 3492  
 Cp 3013 CCAAGCGGCAATTAAGCTTCTACTAAGGAGCAACAAAGCTGGTACCGGGGCCCCCTCG 2954  
 Db 3493 AGGTCGAC 3500  
 Cp 2953 AGGTCGAC 2946

RESULT 9 SYNBLKSPV 2964 bp DNA circular SYN 04-JUN-1993  
 LOCUS

DEFINITION  
 ACCESSION  
 NID  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

BlueScribe KS Plus cloning vector.  
 L08785  
 g310729

Synthetic construct DNA.  
 synthetic construct  
 artificial sequence.  
 1 (bases 1 to 2964)

Gilbert, W.

Obtained from VecBase 3.0

Unpublished (1991)

These data and their annotation were supplied to GenBank by Will Gilbert under the auspices of the GenBank Curator Program.

BlueScribe KS Plus - Cloning vector

ENTRY BLUESKP

TITLE BlueScribe KS Plus - Cloning vector

DATE 28-JAN-1987

#sequence 02-FEB-1987

#sequence 04-MAR-1987

#sequence 03-APR-1987

ACCESSION V00078

SOURCE artificial

REFERENCE

#number 1

#authors Fernandez J.M., Short J.M., Renshaw M., Huse W., Sorae J.

#journal Gene (1987) in press

#citation Sequence data from StrataGene

#comment sequence correction according to StrataGene COMMENT

Obtained from StrataGene on floppy disc.

Revised 02-FEB-1987 by F. Pfeiffer:

1409/10 'Ar' to 'TA' to match revised sequence of pBR322

Revised 4-MAR-1987 to match sequence of pUC19 on request

of StrataGene

Polylinker region revised 03-APR-1987 according to StrataGene

COMMENT

The stand shown corresponds to pUC19c.

As in the published sequence of pUC19c, The M13mp19 lacZ region

is on the complementary strand.

COMMENT

This vector contains the fl origin so that the plus strand

can be obtained upon fl superinfection.

KEYWORDS

CROSSREFERENCE

#parent

VecBase(3):BlueM13p

#parent

VecBase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3,

VecSource(3):bgalks, GenBank(50):PFI

#brother

VecBase(3):BlueKSm, VecBase(3):BlueSkp

PARENT

Features of BlueKsp (2964 bp)

residue source

3- 458 5943-5488 (c) phage fl

460- 524 236- 400 pUC19c

626- 645 1- 20 T7 promoter

653- 760 108- 1 (c) BlueKs-polylinker

772- 791 20- 1 (c) T3 promoter

795-2964 448-2617 pUC19c

Conflict (cfl) and Mutations (mut): none

PARENT

Features of BlueKsp (2964 bp)

residue source

3- 458 5943-5488 (c) phage fl

460- 624 449- 285 (c) pUC19

626- 645 1- 20 T7 promoter

653- 760 108- 1 (c) BlueKs-polylinker

772- 791 20- 1 (c) T3 promoter

795-1031 237- 1 (c) pUC19

1032-2964 2686- 754 (c) pUC19

Conflict (cfl) and Mutations (mut): none

FEATURE











1020 East Meadow Circle, Palo Alto, CA 94303, USA  
This vector can be obtained from CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303, USA. To place an order call (415) 424-8222 or (800) 662-2566, extension 1. International customers, please contact your local distributor. For technical information, call (415) 424-8222 or (800) 662-2566, extension 3.

This sequence has been compiled from information in the sequence databases, published literature and other sources, together with partial sequences obtained by CLONTECH: this vector has not been completely sequenced. If you suspect there is an error in this sequence, please contact CLONTECH's Technical Service Department at (415) 424-8222 or (800) 662-2566, extension 3 or E-mail: [TECH@CLONTECH.COM](mailto:TECH@CLONTECH.COM).

FEATURES	source	Location/Qualifiers
BASE COUNT	726 a 756 c 735 g 750 t	
ORIGIN		
Query Match	43.0%	Score 2224; DB 32; Length 2967;
Best Local Similarity	99.9%	Pred. No. 0.00e+00;
Matches	226; Conservative	0; Mismatches 2; Indels 0; Gaps
		/organism="Cloning vector pDIRECT"
		/db_xref="taxon:31799"
Db	700	GTGGCACTTTTGGGCAAAATGTGGGGAACCCCTATTGTTTATTTTCTAATACATT 759
Cp	5173	GTGGCACTTTTGGGGAATGTGGGGAACCCCTATTGTTTATTTTCTAATACATT 5114
Db	750	CAAAATGATCGGCTCATGAGCAATACCCGTGATAAATGCTTCAATATATTGAAAA 819
Cp	5113	CAAAATGATCGGCTCATGAGCAATACCCGTGATAAATGCTTCAATATATTGAAAA 5054
Db	820	GGAGAGTATGAGTATTCACATTTCCGHTGCGCCCTATTCCCTTTTTGCGGCATTTT 879
Cp	5053	GGAGAGTATGAGTATTCACATTTCCGHTGCGCCCTATTCCCTTTTTGCGGCATTTT 4994
Db	880	GCTTTCCTGTTTTGCTGCACCAGAAAGCTGGTGAAGTAAAGATGCTGAAGTCAGT 939
Cp	4993	GCTTTCCTGTTTTGCTGCACCAGAAAGCTGGTGAAGTAAAGATGCTGAAGTCAGT 4934
Db	940	TGGTGACAGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTT 999
Cp	4933	TGGTGACAGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTT 4874
Db	1000	TTGCGCCCGGAAGAACGTTTTTCGAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGG 1059
Cp	4873	TTGCGCCCGGAAGAACGTTTTTCGAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGG 4814
Db	1060	TATTATCCGCTATTGACGCGGGGCAAGACCACTGGGTGGGGAATACACTATTCTCAGA 1119
Cp	4813	TATTATCCGCTATTGACGCGGGGCAAGACCACTGGGTGGGGAATACACTATTCTCAGA 4754
Db	1120	ATGACTTTGTTGAGTACTACCAAGTACAGAAAAACATCTTACGGATGGCATGACAGTAA 1179
Cp	4753	ATGACTTTGTTGAGTACTACCAAGTACAGAAAAACATCTTACGGATGGCATGACAGTAA 4694
Db	1180	GAGAAATTACGATGCTGCCATAACCATGATGATAACACTTGGCGGCNACTTACTTCTGA 1239
Cp	4593	GAGAAATTACGATGCTGCCATAACCATGATGATAACACTTGGCGGCNACTTACTTCTGA 4634
Db	1240	CAACGATCGGAGGACCGAAGGAGCTTAACGGCTTTTTTGCACAACTGGGGATCATGTAA 1299
Cp	4613	CAACGATCGGAGACCGAAGGAGCTTAACGGCTTTTTTGCACAACTGGGGATCATGTAA 4574
Db	1300	CTCGCTTTGATCGTTGGGAACCGGAGCTGAATTAAGCCATACCAAACGACGAGGTGACA 1359
Cp	4573	CTCGCTTTGATCGTTGGGAACCGGAGCTGAATTAAGCCATACCAAACGACGAGGTGACA 4514
Db	1360	CCACGATCCTCTAGCAATGGCAACAACTTCGCGAAACTATTAACTGGCGGAACACTTA 1419
Cp	4513	CCACGATCCTCTAGCAATGGCAACAACTTCGCGAAACTATTAACTGGCGGAACACTTA 4454



Db	2765	CTCACATGTTCTTCTTCTCGGTTATCCCTGATTCCTGATGATGCGGATTAACCGCTTATACCGCTTTTG	2828			
Cp	3373	CTCACATGTTCTTCTTCTCGGTTATCCCTGATTCCTGATGATGCGGATTAACCGCTTATACCGCTTTTG	3314			
Db	2825	AGTGAGCTCATACCGCTCGCGCAGCGACCGCAAGCAGCGAGCGAGCTCAGTCAGTGAGTGAGG	2884			
Cp	3313	AGTGAGCTCATACCGCTCGCGCAGCGACCGCAAGCAGCGAGCTCAGTCAGTGAGTGAGG	3254			
Db	2885	AAGCGGAAGAGCGCCCAATATACGCAAAACCGCTCTCTCCCGCGCGCTTGGCCGATCATTAAT	2944			
Cp	3253	AAGCGGAAGAGCGCCCAATATACGCAAAACCGCTCTCTCCCGCGCGCTTGGCCGATCATTAAT	3194			
Db	2945	GCAGCTGGCAGCAGCAGGTTTCCCGACTGGAAGCGGGCAGTGACGCGCAACGCAATTAATG	3034			
Cp	3193	GCAGCTGGCAGCAGCAGGTTTCCCGACTGGAAGCGGGCAGTGACGCGCAACGCAATTAATG	3134			
Db	3005	TGAGTTAGCTCACTCAATAGCAGCCCGAGGCTTTACACTTATGCTTCCGSCCTGTAIGT	3064			
Cp	3133	TGAGTTAGCTCACTCAATAGGCAACCCAGGCTTTACACTTATGCTTCCGSCCTGTAIGT	3074			
Db	3065	TGTTGGAAATTGTAGCGGATAAACAATTTACACAGSAAACAGCTATGACCATGATTACG	3124			
Cp	3073	TGTTGGAAATTGTAGCGGATAAACAATTTACACAGSAAACAGCTATGACCATGATTACG	3014			
Db	3125	CCAAGCGCGCAATTAAACCTCACTAAGAGCAACAAAGCTGGGTACCGGGCCCCCGCTCG	3184			
Cp	3013	CCAAGCGCGCAATTAAACCTCACTAAGAGCAACAAAGCTGGGTACCGGGCCCCCGCTCG	2654			
Db	3185	AGGTC 3189				
Cp	2953	AGGTC 2949				
RESULT 13						
LOCUS	A58522	7474 bp	DNA circular PAT 05-MAR-1998			
DEFINITION	Sequence 26 from Patent WO9638555.					
ACCESSION	A58522					
NID	g3714138					
KEYWORDS	unidentified.					
SOURCE	unidentified.					
ORGANISM	unclassified.					
REFERENCE	1 (bases 1 to 7474)					
AUTHORS	Bogert, R., Sringham, E. and Vandekerckhove, J.					
TITLE	PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR USE IN THE CONTROL OF CELL BEHAVIOUR					
JOURNAL	Patent: WO 9638555-A 26 05-DEC-1996;					
COMMENT	BOGAERT THIERRY (BE)					
FEATURES	Other publication AU 6123496 961218.					
source	Location/Qualifiers					
	1..7474					
	/organism="unidentified"					
	/db_xref="taxon:32644"					
BASE COUNT	2107 a	1921 c	1656 g 1790 t			
ORIGIN						
Query Match 43.0%; Score 2222; DB 22. Length 7474.						
Best Local Similarity 100.0%; Pred. No. 0.00e+00;						
Matches 2222; Conservative 0; Mismatches 0; Indels 0; Gaps						
Db	5253	CTCAGGGGGGGGGCCCGTACCAGCTTTTGTTCCCTTTAGTGAGGTTTAATTCGCGCTT	5312			
Qy	2952	CTCAGGGGGGGGGCCCGTACCAGCTTTTGTTCCCTTTAGTGAGGTTTAATTCGCGCTT	3011			
Db	5313	GGCGTAATCATGCTATAGCTGTTCCCTGTGTGAATTTGTAACGGGTACAAATTCACA	5372			
Qy	3012	GGCGTAATCATGCTATAGCTGTTCCCTGTGTGAATTTGTAATTCGCTACAAATTCACA	3071			
Db	5373	CAACATACGAGCGGGAAGCATAAATGTAAAGCCCTGGGTCGCTTAATGAGTGAAGTAACT	5432			
Qy	3072	CAACATACGAGCGGGAAGCATAAATGTAAAGCCCTGGGTCGCTTAATGAGTGAAGTAACT	3131			



## KEYWORDS

SOURCE Cloning vector pFUN.  
 ORGANISM Cloning vector pFUN  
 artificial sequence; cloning vectors.  
 REFERENCE 1 (bases 1 to 8072)  
 AUTHORS Poquet, I., Ehrlich, S.D. and Gruss, A.  
 TITLE An export-specific reporter designed for gram-positive bacteria:  
 application to *Lactococcus lactis*  
 J. Bacteriol. 180 (7): 1904-1912 (1998)  
 JOURNAL 98196737  
 MEDLINE 2 (bases 1 to 8072)  
 REFERENCE Poquet, I. and Gruss, A.  
 AUTHORS Direct Submission  
 TITLE Submitted (15-DEC-1997) Laboratoire de Genetique Appliquee-URLGA,  
 JOURNAL Institut National de la Recherche Agronomique, CRJ, Jouy en Josas  
 78352, France

## FEATURES

Location/Qualifiers  
 1..8072  
 /organism="Cloning vector pFUN"  
 /db\_xref="taxon:74538"  
 /note="plasmid vector that creates translational fusions  
 with a nuclease reporter without a signal peptide"  
 misc\_feature 1..4580  
 /note="pL252 moiety; similar to sequence deposited under  
 GenBank Accession Number AF039139"  
 misc\_feature 4581..7472  
 /note="similar to BlueScript"  
 terminator 7473..7507  
 /note="trpA terminator"  
 misc\_feature 7512..7547  
 /note="multiple cloning site"  
 gene <7543..8010  
 /note="truncated nuc gene potentially encoding deltaSP-Nuc  
 ORF"  
 CDS <7543..8010  
 /gene="delta-nuc"  
 /note="nuclease from *Staphylococcus aureus* devoid of its  
 signal peptide and the thirteen amino acids at the  
 N-terminus of mature protein"  
 /codon\_start=1  
 /transl\_table=11  
 /product="deltaSP-Nuc"  
 /db\_xref="PID:G3043925"  
 /translation="DPTVYSATSKKLHKEPATLIKIDGTVKLMKGPMTFRLL  
 VTPTEKPKKGVKGYGEASFTKKMVENAKTIEVEFDKQRTDKYGRGLAVIADG  
 KWNVALVROGLAKVAVYKPNNTHEQHLRSEAKKEKLNINSEDNADSGQ"  
 BASE COUNT 2717 a 1474 c 1694 g 2187 t

Query Match 43.0% Score 2224; DB 32; Length 8072;  
 Best Local Similarity 99.9% Pred. No. 0.00e+00;  
 Matches 2226; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 5240 GTGGCACTTTTCGGGAAATGTGCGGGAACCCCTATTGTTTATTTTCTAAATACATT 5299  
 |||||  
 Cp 5173 GTGGCACTTTTCGGGAAATGTGCGGGAACCCCTATTGTTTATTTTCTAAATACATT 5114  
 |||||  
 Db 5300 CAATAATGATATCGCTCATGAGACATAACCCCTGATAAATGCTTCAATAATATTGAAAAA 5359  
 |||||  
 Cp 5113 CAATAATGATATCGCTCATGAGACATAACCCCTGATAAATGCTTCAATAATATTGAAAAA 5054  
 |||||  
 Db 5350 GGAAGATGATGATTTCAACATTTCCGTGTGCGCCCTTAITCCCTTTTTTGGGGCATTTT 5419  
 |||||  
 Cp 5053 GGAAGATGATGATTTCAACATTTCCGTGTGCGCCCTTAITCCCTTTTTTGGGGCATTTT 4994  
 |||||  
 Db 5420 GCTTCTCTGTTTTGTCTACCCGAGAACGCTGTGGAAGTAAAGATGCTGAAGATCAGT 5479  
 |||||  
 Cp 4993 GCTTCTCTGTTTTGTCTACCCGAGAACGCTGTGGAAGTAAAGATGCTGAAGATCAGT 4934  
 |||||  
 Db 5480 TGGGTGCACGAGTGGGTGTACATCGAACTGGATCTCAACAGCGGTGAAGATCCCTTGAGAGTT 5539  
 |||||  
 Cp 4933 TGGGTGCACGAGTGGGTGTACATCGAACTGGATCTCAACAGCGGTGAAGATCCCTTGAGAGTT 4874

Db 5540 TTGCGCCGGAAGAACGTTTTTCAATGATGATGACACTTTTAAAGTTTCTGCTATGTTGCGCGG 5599  
 |||||  
 Cp 4873 TTGCGCCGGAAGAACGTTTTTCAATGATGATGACACTTTTAAAGTTTCTGCTATGTTGCGCGG 4814  
 |||||  
 Db 5600 TATTATCCCGTATTGACGCCGCGGAAGAACAACTGCGTGGCGCATACATAATTCTCAGA 5659  
 |||||  
 Cp 4813 TATTATCCCGTATTGACGCCGCGGAAGAACAACTGCGTGGCGCATACATAATTCTCAGA 4754  
 |||||  
 Db 5660 ATGACTTGGTTGACTACTCACCAGTCACAGAAAGCATCTTTACGGATGGCATGACATAA 5719  
 |||||  
 Cp 4753 ATGACTTGGTTGACTACTCACCAGTCACAGAAAGCATCTTTACGGATGGCATGACATAA 4694  
 |||||  
 Db 5720 GAGAATATGCACTGCTGTCATAACCATGATGATAACACTGCGGCCAACTTACTTTTGA 5779  
 |||||  
 Cp 4693 GAGAATATGCACTGCTGTCATAACCATGATGATAACACTGCGGCCAACTTACTTTTGA 4634  
 |||||  
 Db 5780 CAACGATCGGAGCGGAGGAGCTAACCCGTTTTTTGCACAACTATTAACCTGGCGCAACTCTTA 5839  
 |||||  
 Cp 4633 CAACGATCGGAGCGGAGGAGCTAACCCGTTTTTTGCACAACTATTAACCTGGCGCAACTCTTA 4574  
 |||||  
 Db 5840 CTCGCTTCATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAGACAGAGGSIACAA 5899  
 |||||  
 Cp 4573 CTCGCTTCATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAGACAGAGGSIACAA 4514  
 |||||  
 Db 5900 CCACGATGCTGTAGCAATGCAACAACGTTGCGCAAACTATTAACTGGCGCAACTCTTA 5959  
 |||||  
 Cp 4513 CCACGATGCTGTAGCAATGCAACAACGTTGCGCAAACTATTAACTGGCGCAACTCTTA 4454  
 |||||  
 Db 5960 CTCAGCTCCCGGCAACAATTAATAGACAGTGGTGGAGGAGGAGTAAAGTTGACAGAGACAC 6019  
 |||||  
 Cp 4453 CTCAGCTCCCGGCAACAATTAATAGACAGTGGTGGAGGAGGAGTAAAGTTGACAGAGACAC 4394  
 |||||  
 Db 6020 TTCTGCGCTCGGCCCTTCGCGTGGCTGTTTATTCSTSAIAAAATCTGSAAGGCTGAGC 6079  
 |||||  
 Cp 4393 TTCTGCGCTCGGCCCTTCGCGTGGCTGTTTATTCSTSAIAAAATCTGSAAGGCTGAGC 4334  
 |||||  
 Db 6080 GTGGGTCTCGCGCTATCATTGCGACACTGGGCGCAAGATGGTAAAGGCTCCCGGATCTGAG 6149  
 |||||  
 Cp 4333 GTGGGTCTCGCGCTATCATTGCGACACTGGGCGCAAGATGGTAAAGGCTCCCGGATCTGAG 4274  
 |||||  
 Db 6140 TTACTACACGACGGGGAGTCAGGCAACTATGGATGAACAAATAGACAGATCGGCIACAA 6199  
 |||||  
 Cp 4273 TTACTACACGACGGGGAGTCAGGCAACTATGGATGAACAAATAGACAGATCGGCIACAA 4214  
 |||||  
 Db 6200 TAGTGCCTCAGCTGATTAAGCATTCGTAACTGTCAGACCAAGTTTACTCATATATCTTT 6259  
 |||||  
 Cp 4213 TAGTGCCTCAGCTGATTAAGCATTCGTAACTGTCAGACCAAGTTTACTCATATATCTTT 4154  
 |||||  
 Db 6260 AGATTGATTTAAAACTTCATTTTAAATTAAGGATCTAGGTGAAGATCGCTTTTGTATA 6319  
 |||||  
 Cp 4153 AGATTGATTTAAAACTTCATTTTAAATTAAGGATCTAGGTGAAGATCGCTTTTGTATA 4094  
 |||||  
 Db 6320 ATCTCATGACCAAAATCCCTTAACGTGAGTTTTCCTTCGACTGAGCTGACACACCGCTAG 6379  
 |||||  
 Cp 4093 ATCTCATGACCAAAATCCCTTAACGTGAGTTTTCCTTCGACTGAGCTGACACACCGCTAG 4034  
 |||||  
 Db 6380 AAAAGATCAAGAGATCTCTCTGAGATCCCTTTTCTCGGGTAACTCTGCTGCTTGCAA 6439  
 |||||  
 Cp 4033 AAAAGATCAAGAGATCTCTCTGAGATCCCTTTTCTCGGGTAACTCTGCTGCTTGCAA 3974  
 |||||  
 Db 6440 CAAAAAACCACCGCTTACCAGCGGTGGTTTGTTCGCGGATCAAGAGCTACCAACTCTTT 6499  
 |||||  
 Cp 3973 CAAAAAACCACCGCTTACCAGCGGTGGTTTGTTCGCGGATCAAGAGCTACCAACTCTTT 3914  
 |||||  
 Db 6500 TTCGGAAGTAACTGGCTTCACGAGCGGACATACCAAACTACTGCTTCTAGTGTAGC 6559  
 |||||  
 Cp 3913 TTCGGAAGTAACTGGCTTCACGAGCGGACATACCAAACTACTGCTTCTAGTGTAGC 3854  
 |||||  
 Db 6560 CGTAGTTAGCGCACCACTCAAGAACTCTGTAGACCGGCTACATACCTCGCTCTGTCTAA 6619  
 |||||  
 Cp 3853 CGTAGTTAGCGCACCACTCAAGAACTCTGTAGACCGGCTACATACCTCGCTCTGTCTAA 3754



Db	2534	CAAGTCATCTTGAGAACTAGTATGCGGCGACCGAGTTGCTCTTTGCGCGGGGTCAATACG	2593
Qy	4746	CAAGTCATCTTGAGAACTAGTATGCGGCGACCGAGTTGCTCTTTGCGCGGGGTCAATACG	4805
Db	2594	GGATAATACGGCGCCACATAGCAGAACTTTAAAAGTGCTCAICATTTGAAAAAGCTTTTC	2653
Qy	4806	GGATAATACGGCGCCACATAGCAGAACTTTAAAAGTGCTCAICATTTGAAAAAGCTTTTC	4865
Db	2654	GGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTCG	2713
Qy	4866	GGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTCG	4925
Db	2714	TGACCCCAACTGATCTTCAGCATCTTTTACTTTTCACCAAGCGTTTCTGGTGAGCAAAAC	2773
Qy	4926	TGACCCCAACTGATCTTCAGCATCTTTTACTTTTCACCAAGCGTTTCTGGTGAGCAAAAC	4985
Db	2774	AGGAAGCAAAATTCGCCAAAAAAGGGAATAAGGCGCACACGGAATGTTCAATACTCAT	2833
Qy	4986	AGGAAGCAAAATTCGCCAAAAAAGGGAATAAGGCGCACACGGAATGTTCAATACTCAT	5045
Db	2834	ACTCTCCCTTTTCAATATATTGAAGCATTTATCAGGGTTATTGTCTCATGACGGATA	2893
Qy	5046	ACTCTCCCTTTTCAATATATTGAAGCATTTATCAGGGTTATTGTCTCATGACGGATA	5105
Db	2894	CATATTGAAATGATTATAGAAAAATAACAAATAGGGGTTCCGGGCACATTTCCCGGAAA	2953
Qy	5106	CATATTGAAATGATTATAGAAAAATAACAAATAGGGGTTCCGGGCACATTTCCCGGAAA	5165
Db	2954	AGTGC 2958	
Qy	5166	AGTGC 5170	

Search completed: Tue Mar 16 09:48:54 1999  
Job time : 13754 secs.





\*\*\*\*\*  
W P E R L  
\*\*\*\*\*  
(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_n n a - n a database search, using Smith-Waterman algorithm  
Run on: Tue Mar 16 12:05:21 1999; MasPar time 731 94 Seconds  
Tabular output not generated. 961.559 Million cell updates/sec

Title: >US-09-020-716-7  
Description: (1-5173) from US09020716.seq  
Perfect Score: 5173  
N.A. Sequence: 1 CTAATTTGTAAGCGTTAATA... ATTCCCGGAAAGTGCCAC 5173  
Comp GATTTAACATTCGAATTAT... TAAAGGGGCTTTTCACGNTG

Scoring table: TABLE default  
Gap 5

Nmatch STD : Dbase 0; Query 0

Searched: 189442 seqs, 58026449 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-genes32  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39 40:part40

Statistics Mean 10.428, Variance 6.781, scale 1.538

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	3222	43.0	5178	27	149876	0.00e+00
2	3222	43.0	7474	31	771320	0.00e+00
3	3220	42.9	6202	27	149877	0.00e+00
4	3203	42.6	4145	7	Q40279	0.00e+00
5	3203	42.6	5314	2	U13576	0.00e+00
6	3202	42.6	5356	25	T43794	0.00e+00
7	3202	42.6	5534	25	T43137	0.00e+00
8	3201	42.5	3681	2	Q13378	0.00e+00
9	3198	42.5	3699	39	V14340	0.00e+00
10	3198	42.5	4277	7	Q40280	0.00e+00
11	3196	42.5	4539	14	Q87347	0.00e+00
12	3200	42.5	4792	15	Q84596	0.00e+00
13	3200	42.5	5042	15	Q84594	0.00e+00

C	14	2198	42.5	6226	7	Q40281	Sequence of clone pV2	0.00e+00
C	15	2198	42.5	7287	39	V02042	Plasmid pWRG3169 enco	0.00e+00
C	16	2196	42.5	10930	20	O81226	Plasmid pM16-1.	0.00e+00
C	17	2196	42.5	10950	20	O81225	Plasmid pM16.	0.00e+00
C	18	2195	42.4	14311	24	T38744	Plasmid pVSVFL(+)-inc	0.00e+00
C	19	2190	42.3	6824	6	Q39050	K lactis/S cerevisae	0.00e+00
C	20	2185	42.2	7228	9	Q50201	Vector p37M1-10D.	0.00e+00
C	21	2166	41.9	4118	35	T69189	Construct pGEM-hTR(Ka	0.00e+00
C	22	2166	41.9	4118	35	T69188	Construct pGEM-hTR co	0.00e+00
C	23	2165	41.9	5585	14	O87418	Plasmid pM26-IL2.	0.00e+00
C	24	2167	41.9	6295	39	V02043	Plasmid pWRG3196 enco	0.00e+00
C	25	2162	41.8	4454	30	T66462	Plasmid pCRM1.8 conta	0.00e+00
C	26	2162	41.8	4540	32	T69891	Human cyclin D1-human	0.00e+00
C	27	2162	41.8	4824	20	T08975	Shortened C1 gene in	0.00e+00
C	28	2163	41.8	4883	4	Q25709	PDE110.	0.00e+00
C	29	2163	41.8	4983	9	O51122	PDE110; Plasmid DNA r	0.00e+00
C	30	2162	41.8	5399	4	Q25706	PDE108.	0.00e+00
C	31	2163	41.8	5585	29	T61430	Interleukin-2 express	0.00e+00
C	32	2163	41.8	5620	4	Q27489	PVE108.	0.00e+00
C	33	2163	41.8	5620	9	O53875	Plasmid pVE108 used i	0.00e+00
C	34	2163	41.8	5642	8	O51193	PDE4; plasmid DNA rep	0.00e+00
C	35	2162	41.8	6274	7	O43812	PRLD1D203 Apal const	0.00e+00
C	36	2162	41.8	6555	9	Q53874	Plasmid pVE144 used i	0.00e+00
C	37	2162	41.8	6926	18	T12662	IL-4.Y124D/Igcl fusio	0.00e+00
C	38	2163	41.8	7050	7	Q40419	Sequence of pTE4.	0.00e+00
C	39	2162	41.8	7106	8	Q47193	Plasmid pHCV-167 codi	0.00e+00
C	40	2162	41.8	7106	15	Q47404	Plasmid pHCV-162 codi	0.00e+00
C	41	2162	41.8	7298	8	Q47142	Plasmid pCB51 encodin	0.00e+00
C	42	2163	41.8	10288	31	T71322	Plasmid pB51 encodin	0.00e+00
C	43	2162	41.8	10443	31	T71319	Plasmid pB52 encodin	0.00e+00
C	44	2163	41.8	13414	31	T71321	Plasmid pCB50 encodin	0.00e+00
C	45	2162	41.8	16656	21	T30787	Alphavirus-based euka	0.00e+00

ALIGNMENTS

RESULT 1  
ID T49876 standard: DNA; 5178 BP.  
AC T49876;  
DT 28-MAR-1997 (first entry)  
DE Pret-Splice.  
KW Pret-Splice; pTet-trak; tetracycline transactivator; promoter;  
KW gene expression; transgenic animal; animal model; drug screening;  
KW vector, ds.  
CS Synthetic.  
FH Key  
FT misc\_rna  
FT Location/Qualifiers  
FT complement (2591)  
FT /tag= a  
FT /note= "putative start site of trxn"  
FT complement (2615..2622)  
FT /tag= b  
FT WO9640946-A1.  
PN 19-DEC-1996.  
PD 07-JUN-1996; U10109.  
PF 07-JUN-1996; US-474169.  
PR (OYVA ) UNIV YALE.  
PA Schatz DG;  
PI WPI; 97-077273/07.  
DR Nucleic acid encoding tetracycline transactivator fusion protein -  
PT provides rapid and reversible control of gene expression, e.g. for  
PT creating animal models for drug screening  
PS Example 1; Fig 9b-g; 82pp; English.  
CC Vector pret-Splice (T49876) is utilised in the construction of  
CC autoregulatory vector plasmid pret-trak (149877). Splice-PA was  
CC made by ligating the SV40 small T antigen intervening sequence and  
CC the SV40 early polyA sequence into pBSKII+. The XhoI-SalI fragment  
CC of pUC13-3 (contg. 7 copies of the tet operator upstream of minimal  
CC promoter retp) was cloned upstream of the splice/polyA sequence of  
CC pSplice-PA to form pret-Splice. A modified tetracycline  
CC transactivator (trak) gene was cloned into pret-Splice to form pret-  
CC trak. The construct provides rapid, reversible control of gene  
CC expression in eukaryotic cells or transgenic animals, e.g. for  
CC heterologous protein, RNA or antisense sequence prodn., or for





[illegible]

11153

REF: T49877 standard: DNA: 6206 BP.

AC T49877;

DT 28-MAR-1997 (first entry)

DE Anteroviral+ery vector plasmid pTet+TAK.

kw pTet::trkA; pTet::Splice; tetacycline transactivator, promoter;

KW. *quercus velutina*

KW : vector: ds

[illegible]

Db	5665	tgccgaagaatgaagtggccgcagtgattatcactcatgagttatgacagcacatgcataattc	5724
Qy	4632	TGTCAGAAAGTAAGTTGGCGCAGTGTATCACTCAATGTTATGGCAGCACTGCATAATTC	4691
Db	5725	tcttactgtcaatccacgtcaagatgctttctgtcaactagtagtactcaaccaatc	5784
Qy	4692	TCATTACTGTGATCCGATCGTAAAGATGCTTTCTGTGACTGTGTAGTACTCAACCAAGTC	4751
Db	5785	attctdagaatagtgtatgcggcgaccaggttgccttgcgcggcgctcaatcacgggataa	5844
Qy	4752	ATTCTGAGAAATAGTGTATGGCGGACCGAGTTGCTCTTTGCCCGGCGTCAATACGGGGATAA	4811
Db	5845	taccggcgccacatagcagaacctttaaaadtgcctacatcttgaaaaaacgtttctcggggcg	5904
Qy	4812	TACCGGCCACATATAGCAGAACITTTAAAAGTGTCTCATCTATGGAAGAGCTTCTTCGGGGCG	4871
Db	5905	aaactctcaagatcttaccctcttaagatccagttcagatgaaccacctcatgcacc	5964
Qy	4872	AAACTCTTAAGGATCTTACCGGCTGTGTGAGATCTAGTTCGATGAACCCACTCTGTGCACC	4931
Db	5965	caactgattctcagcatctttactttcacccagcgtttctgggtgagcaaaaacaggaag	6024
Qy	4932	CAACTGATCTTCAGCATCTTTACITTCACCAAGCGTTTCTGGGTGAGTAAAAACAGGAG	4991
Db	6025	gcataatgcgcgaaaaaagggaataagggcgacacgggaaatgtttaaatactcatactctt	6084
Qy	4992	GCAAAATGCGCGAAAAAAGGGAAATAGGCGGACACGGAAATGTTGAATACTCAATCTT	5051
Db	6085	cccttttcaatattatgaagcatttatcagggttatgtctcatgacgcgatacatatt	6144
Qy	5052	CCCTTTTCAAATATTGAAGCAATTAACAGGATTAATGCTCATGAGCGATACATATT	5111
Db	6145	tgaatgtatttgaataataaacaataaggggttcgcgcgcacattccccgaaaaagtcgc	6204
Qy	5112	TGAATGTATTACAAAATTAACAAATAGGGGTTCCGCGGCACATTTCCCCGAAAGTGCC	5171
Db	6205	ac 6206	
Qy	5172	ac 5173	

```

RESULT      4
ID          Q40279 standard; DNA: 4145 BP.
AC          Q40279;
DT          02- AUG-1993 (first entry)
DE          Sequence of clone pS2qpt-S4.
KW          Plasmid; cloning; restriction site; ss.
OS          Synthetic.
FH          Key
FT          misc_feature
FT          1..2225
FT          Location/Qualifiers
FT          /tag= a
FT          /label= pN2qpt-S4
FT          /note= "Posn. 1 corresp. to the first nucleotide
FT          G '5-TGGCACTTT TCGGGGAAAT-3'"
FT          misc_feature
FT          2227..2236
FT          /tag= b
FT          /label= SmaI adaptor
FT          /label= SmaI adaptor
FT          2396..2851
FT          /tag= c
FT          /label= E. coli gpt gene
FT          complement (3081..3323)
FT          /tag= d
FT          /label= vaccinia p7.5 promoter
FT          3358..3451
FT          /tag= e
FT          /label= S4 of pN2qpt-S4
FT          /note= "Cligo P-attP(9)"
FT          2237..4145
FT          /tag= f
FT          /label= pN2qpt-S4
FT          AU9221269-A.
FT          04-MAR-1993.
FT          25-AUG-1993.

```

















Db 2712 gaacatattcagagatctatatttggttcacatatttaccctgactgcccgtcg 2771  
 QY 4207 ccaactatcagagatctatatttggttcacatatttaccctgactgcccgtcg 4266  
 Db 2772 tagataactcagatcagagagagagagagagagagagagagagagagagag 2831  
 QY 4267 tagataactcagatcagagagagagagagagagagagagagagagagagag 4326  
 Db 2832 gaacacagctcagagctcagagctcagagctcagagctcagagctcagagctc 2891  
 QY 4327 gacacacagctcagagctcagagctcagagctcagagctcagagctcagagctc 4386  
 Db 2892 gacacacagctcagagctcagagctcagagctcagagctcagagctcagagctc 2951  
 QY 4387 gacacacagctcagagctcagagctcagagctcagagctcagagctcagagctc 4446  
 Db 2952 actagagtaadtagctcagagctcagagctcagagctcagagctcagagctc 3011  
 QY 4447 actagagtaadtagctcagagctcagagctcagagctcagagctcagagctc 4506  
 Db 3012 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 3071  
 QY 4507 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 4566  
 Db 3072 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 3131  
 QY 4567 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 4626  
 Db 3132 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 3191  
 QY 4627 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 4686  
 Db 3192 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 3251  
 QY 4687 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 4746  
 Db 3252 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 3311  
 QY 4747 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 4806  
 Db 3312 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 3371  
 QY 4807 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 4866  
 Db 3372 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 3431  
 QY 4867 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 4926  
 Db 3432 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 3491  
 QY 4927 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 4986  
 Db 3492 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 3551  
 QY 4987 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 5046  
 Db 3552 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 3611  
 QY 5047 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 5106  
 Db 3612 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 3671  
 QY 5107 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 5166  
 Db 3672 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 3731  
 QY 5167 atcgtggttcacagctcgtctggttggatggttggatggttggatggttggatg 5230

RESULT 4  
 ID V14340 standard: DNA: 3699 BP.  
 AC V14340;

DT 19-MAY-1998 (first entry)  
 DE Plasmid pBSGFP expressing green fluorescent protein.  
 KW Green fluorescent protein; GFP; blue fluorescent protein; BFP; jellyfish  
 KW genetic engineering marker; gene therapy; plasmid; ss.  
 OS Synthetic.  
 PN W09742320-A1.  
 PD 13-NOV-1997.  
 PF 07-MAY-1997; 007625.  
 PR 08-MAY-1996; US-646538.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Gattaris CA, Pavlakis GN, Stauber RH, Vournakis JN.  
 DR WPI: 97-558982/51.  
 PT New nucleic acid encoding proteins of Aequorea victoria with  
 PT increased fluorescence useful as markers for detecting cellular  
 PT transformation, subcellular localisation of proteins, for assessing  
 PT gene therapy, mutagenicity etc.  
 PS Example -; Page 67-68; 105pp; English.  
 CC This sequence is a plasmid used to express the mutated green fluorescent  
 CC protein (GFP) of Aequorea victoria coding sequence of the invention. The  
 CC DNA of the invention that encodes a protein that is (a) a protein that  
 CC has Leu at position 65 (and optionally Thr at 168 and optionally further  
 CC Cys at 66) and has cellular fluorescence at least 5 times that of  
 CC wild-type GFP or (b) is a blue fluorescent protein (BFP) with His at  
 CC position 67 and also at least one of Leu at 65 and Ala at 164, and has  
 CC cellular fluorescence at least 5 times that of BFP (Tyr67 to His). The  
 CC nucleic acids can be used as markers in genetic engineering and gene  
 CC therapy. They may also be used to detect and characterise regulatory and  
 CC coding sequence elements that control subcellular expression and  
 CC targeting of proteins. Typical applications are monitoring targeting and  
 CC transport of proteins in cells; assessment of gene therapy procedures; in  
 CC diagnosis (when expressed under control of a promoter induced by a  
 CC particular analyte); assessment of mutagenicity of compounds; and for  
 CC drug screening (where expression is controlled by the promoter of a  
 CC target gene), particularly for antiviral or antiparasitic agents.  
 SQ Sequence 3699 BP; 942 A, 884 C, 898 G, 975 T;

Query Match 42.5%; Score 2198; DB 39; Length 4699;  
 Best Local Similarity 99.9%; Pred. No. 0.00+00;  
 Matches 2200; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Db 1495 ccagctttgtccctttagtgagggtaattccagagcttgcctgactgactgactg 1554  
 QY 2572 ccagctttgtccctttagtgagggtaattccagagcttgcctgactgactgactg 4041  
 Db 1555 tgttccgtgtgaaattgttaccctcacatccacacacacacacacacacacacac 1614  
 QY 3032 tgttccgtgtgaaattgttaccctcacatccacacacacacacacacacacacac 4091  
 Db 1615 taaagtataaagcctggagctgactgactgactgactgactgactgactgactgact 1674  
 QY 3092 taaagtataaagcctggagctgactgactgactgactgactgactgactgactgact 4161  
 Db 1675 cactgagcgtcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 1744  
 QY 3152 cactgagcgtcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 4211  
 Db 1735 gcgcggggagagcgggtttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 1794  
 QY 3212 gcgcggggagagcgggtttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 4271  
 Db 1795 tgcgtcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 1854  
 QY 3272 tgcgtcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 4331  
 Db 1855 tatccacagaatcaggagataaagcagagagagagagagagagagagagagagagagag 1914  
 QY 3332 tatccacagaatcaggagataaagcagagagagagagagagagagagagagagagagag 4391  
 Db 1915 ccagggaacccataaaagccgcgttactgactgactgactgactgactgactgactgactg 1974  
 QY 3392 ccagggaacccataaaagccgcgttactgactgactgactgactgactgactgactgactg 4451

Db 1975 agcatcaaaaatgacgctcaagtcaagtcagagtgaggcgaaacccgacagagactataaagat 2034  
 QY 3452 AGCATCAAAAATGACGCTCAAGTCAGAGGTGGGAAACCGGACAGGACTATAAGAT 3511  
 Db 2035 accagcggttccccctgaaagctccctctgctgctctctctctctctctctctctctct 2094  
 QY 3512 ACCAGCGGTTCCCCCTGAAAGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3571  
 Db 2095 ccgatacctgtccgcttctccctctcgggaagcgctgagcgcttctctctctctctctct 2154  
 QY 3572 CCGGATACCTGTCCGCTTCTCTCTCTGCGGAAGCGTGGCGCTTCTCTATAGCTACAGCT 3631  
 Db 2155 gtaagtatctcaagtcgggtgaagtcgctcccaagctgggtgtgtgtgtgtgtgtgtgtgt 2214  
 QY 3632 GTAGTATCTCACTTCGGGTGAGTGTGCTGCTCCAGCTGGGCTGTGTGCACGAACCCC 3691  
 Db 2215 cccttcagccgaacccctcgccttatccggtaactatcgtcttgagtcceaacccggttaa 2274  
 QY 3692 CCGTTCAGGCGTAACGCTGCGCTTATCGGTAAGCGTGGCGCTTCTCTATAGCTACAGCT 3751  
 Db 2275 qacacacttatccgcaactggcgagcagccactggtaacagagattagcagagcgaggtatg 2334  
 QY 3752 GACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATG 3811  
 Db 2335 taagcgtgtctacagagttcttgaagtggtggccttaactaagctacactagaaggacag 2394  
 QY 3812 TAGCGGTGTCTAAGAGTCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAAGGACAG 3871  
 Db 2395 tatttgatctcgcctctgctgaagcgaagttaccttcggaaaaagagatttgtagctctt 2454  
 QY 3872 TATTGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3931  
 Db 2455 qatccggcaaaaaccccgctggttagcggtggttttttttttttttttttttttttttttt 2514  
 QY 3932 GATCGGTAANAACAAACAGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3991  
 Db 2515 cggcgaaaaaagaagatctcaaaaagatccttttgaatctctctctctctctctctctctct 2574  
 QY 3992 CCGCGCAAAAAGGATCTCAAGAGATCTTGAATCTTGAATCTTGAATCTTGAATCTTGA 4051  
 Db 2575 agtggaaacaaaactcacgttaaggatttgggtcagaagattatcaaaaagagattctca 2634  
 QY 4052 AGTGGAAAGAAAACCTCAGGTTAGGGATTGGTATGAGATTATCAAAAAGGATCTTCA 4111  
 Db 2635 cctagatcccttttaaatataaaatgaagttttaaatacactcaaatcaatgatatatgataaa 2694  
 QY 4112 CCTAGATCTCTTTAAATTAATAATGAAGTTTAAATCAATCTAAGTATATATGATGATA 4171  
 Db 2695 ctgggtctgacagttaccgaatgcttaactcaagtcagggcaactatctcagcgatctgtctat 2754  
 QY 4172 CTGGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTAT 4231  
 Db 2755 ttcgttcacatagttgctgaactcccgctcgtgtagataactacgatacggaggct 2814  
 QY 4232 TTGGTCTATCCATAGTTGCTGACTCCCGCTGCTGTAGATAACTACGATACGGAGGGCT 4291  
 Db 2815 taccatctggcccaagtgctgaatatacccgagaccacgcctcaccggtctccagatt 2874  
 QY 4292 TACCATCTGGCTCCAGTGTGCAATATACCGCGAGACCGACCGCTCCAGATT 4351  
 Db 2875 tatcaacaataaaacagccagccggaagggccgagcgagagaagtggtctgcaactttat 2934  
 QY 4352 TATCAGCAATAAACACAGCCAGCGGAAGGGCGGAGAGAGTGGTCTGCAACTTTAT 4411  
 Db 2935 ccgctccatccagctctattaattgttcggggaagctagagtaagtagtgcagttta 2994  
 QY 4412 CCGCTCCATCCAGTCTATTATTGTGTGCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTA 4471  
 Db 2995 atagtttgcgaacggttggccatctgtacagggcaatcgtgtgtcagcgtcgtcgtttg 3054  
 QY 4472 ATAGTTTGCGAACAGGTTGTCGCAATTGCTACAGGCATCGTGGTGTACGCTCGTGGTTG 4531  
 Db 3055 gtagaggtctattcagctcgggttcccaacgatacgaaggcgaattacatgatcccccatgt 3114

QY 4532 GTATGGCTTCACTTACGCTCCGCTTCCCAACGATCAAGGAGTTACATGATCCCCCATGT 4591  
 Db 3115 tctgcaaaaaagcgcttagctccctccgctccgagtcggttctcagaagaagtgtgccc 3174  
 QY 4592 TGTGCAAAAAGCGGTAGCTCCTCGGTCTCGGATCGTGTGTCAGAAATGAGTTGCGG 4651  
 Db 3175 cagtggtatcactcactggttatggcagcactgcataatctcttactgtcatgccatccg 3234  
 QY 4652 CAGTGTATCACTCATGTTATGCGCAGCAGCTGCATATTCCTTACTGTCTATGGCATCCG 4711  
 Db 3235 taagatgcttttctgactggtgactcaaccaagctcattctcgagaatagtgatgc 3294  
 QY 4712 TAAGATGCTTTTCTGTACTGTGAGTAGTCAACCAAGTCAITCTGAGAAATAGTGTATGC 4771  
 Db 3295 ggcgaccgagttgctcttgcgcggcgctcaatacaggggataatacaccgcccacatacgaag 3354  
 QY 4772 GCGCAGCGAGTTGCTCTTCCCGCGGCTCAATACGGGATAATACCGCCCATAGCAGAA 4831  
 Db 3355 ctttaaaagtgtcactcatttgaaaaaacttctcggggcgaataactctcaggaattctac 3414  
 QY 4832 CTTTAAAGTGTCTCATCTTGGAAACGCTTCTTGGGGCGGAAACTCTCAAGGATCTTAC 4891  
 Db 3415 cgctgttagagatccaggttcgatgtaacccactcgtgcacccaactcgtcttcagcatctt 3474  
 QY 4892 CGCTGTGAGATCCAGTTCGATTAACCCACTGTCGACCCAACTGATCTTCAGCATCTT 4951  
 Db 3475 ttactttcaccagccttctcgggtgaacaaaaacaggaaggaacaaatcccgcaaaaaag 3534  
 QY 4952 TTACTTTACCAGCTCTTCTTGGGTGAGCAAAAACAGGAAGCAAAATGCGGCAAAAAGG 5011  
 Db 3535 gaataaggcgacacggaatgtgataactactactctctctctctctctctctctctctct 3594  
 QY 5012 GAATAAGCGGACACGGAATGTTGAATACTCATACITCTCTCTTTTCAATATTATGAA 5071  
 Db 3595 gcatttaccaggggttattgtctcagtcagcgagatacatttgaatgtatttagaataata 3654  
 QY 5072 GCATTTATCAGGGTATTGTCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATA 5131  
 Db 3655 aacaaatagggttcgcgcacatttcccgaaaaagtgccac 3696  
 QY 5132 AACAAATAGGGTTCGCGCACATTTTCCCGGAAAAGTGCCAC 5173

RESULT 10  
 ID Q4280 standard; DNA; 4277 BP.  
 AC Q4280:  
 DT 02-AUG-1993 (first entry)  
 DE Sequence of clone pS2qpt-P2.  
 KW Plasmid; cloning; restriction site: ss.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_feature 1..3357  
 FT /tag= a  
 FT /label= pN2qpt-S4  
 FT /c= 2396..2851  
 FT /tag= b  
 FT /label= E. coli qpt gene  
 FT complement (3081..3323)  
 FT /tag= c  
 FT /label= vaccinia p7.5 promoter  
 FT /c= 3358..3526  
 FT /tag= d  
 FT /label= P2 according to EP application Avipox  
 FT intergenic region  
 FT /note= "oligo P-artp(9)"  
 FT misc\_feature 3527..4277  
 FT /tag= e  
 FT /label= pN2qpt-S4  
 FT AU9221269-A:  
 PN 04-MAR-1993.  
 PF 25-AUG-1992; 021369.  
 PR 26-AUG-1991; US-750080.



```

Query Match      42.5%; Score 2196; DB 14; Length 4539;
Best Local Similarity 99.9%; Prcd. No. 0.00e+00;
Matches 2198; Conservative 0; Mismatches 2; Indels 0; Gaps 0

Db 2268 agctttgtccctttagtagaggttaatttcgacttgcgttaatcatgtcatagctg 2327
QY 2974 AGCTTTGTCCCTTTAGTAGAGGTTAATTGCGCCTTGGCGGTAAATCATGGTCATAGCTG 3033

```

QY 4054 TGGAACGAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACC 4113











```

|||||
QY 5052 CTTTTTCAATATTATTGAAGCAITTTATCAGGGTTATTGTCTCATGAGCGATACATATT 5111
Db 4978 tgaatatttggaaataaacaataagggttcgcgcacatttccccgaaagtgc 5037
QY 5112 TGAATGATTAGAAAATAAACAATAAGGGTTCGCGGCACATTTCCCGAAAAGTGCC 5171
Db 5038 ac 5039
QY 5172 AC 5173

RESULT 14
ID Q40281 standard: DNA: 6926 BP.
AC Q40281:
DE Q2-AUG-1993 (first entry)
DE Sequence of clone pP2-gp160MN.
KW Plasmid; cloning; restriction site; HIV-1; MN-isolate; ss.
OS Synthetic.
FH Key
FT misc_feature 1..3529 Location/Qualifiers
FT /tag= a
FT /label= pS2gpt-P2 sequences
FT 2396..2851
FT cds
FT /tag= b
FT /label= E. coli gpt gene
FT complement (3081..3323)
FT /tag= c
FT /label= vaccinia P7.5 promoter
FT 2396..2851
FT /tag= d
FT /label= P2 according to EP application Avipox
FT "intergenic region"
FT /note= "oligo P-artp(9)"
FT 6173..6926
FT misc_feature
FT /tag= e
FT /label= pN2gpt-S4
FT 3534..6001
FT cds
FT /tag= f
FT /label= HIV-1 strain MN gp160 sequence (EMBL to
REH1VMNC)
PN A09231269-A.
PD 04-MAR-1993.
PF 25-AUG-1992: 021269.
PR 26-AUG-1991: US-750080.
PR 20-JUL-1992: US-914738.
PA (TMMO ) IMMUNO AG.
PI Dörner F, Falkner FG, Pfeleiderer M, Scheiflinger F;
OR WPL: 93-126461/16.
PT Modified eukaryotic cytoplasmic DNA virus prodn. - involves
PT direct molecular cloning of modified DNA molecule contg.
PT cytoplasmic DNA virus genome
PS Example: Pages 177-180; 206pp; English.
CC Plasmids pN2gpt-S3A and pN2gpt-S4 comprise expression cassettes
CC with a selective marker. These plasmids were constructed by first
CC making plasmids pN2-gpta and pN2-gptb which contain an E. coli gpt
CC gene driven by the vaccinia virus P7.5 promoter, flanked by
CC several unique restriction sites including NotI. A new SmaI site
CC was introduced into the plasmid pN2gpt-S4 resulting in the plasmid
CC pS2gpt-S4. Subsequently the S4-promoter was exchanged by the P2-
CC promoter resulting in the plasmid pS2gpt-P2. For the modification
CC of the gp160-gene, a PCR-generated proximal fragment was exchanged
CC leading to a gp160-gene cassette with a minimal 5'-UTR.
SQ Sequence 6926 BP: 2064 A: 1470 C: 1686 G: 1706 T:
Query Match 42.5% Score 2198; DB 7; Length 6926;
Best Local Similarity 99.9% Pred No 0 00e-00;
Matches 2300; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 1 gtgcacatttccggggaataatgctgcgcgaacccctatttatttttctaaataacatt 60
Cp 5173 GTGCGACATTTTCGGGGAATGTGCGGGAACCCCTATTGTGTTTATTTTCTAAATACATT 5114

```

```

Db 61 caaatatgtatccgctcatgagacaataaacctataaactcttcaataatattgaaa 120
Cp 5113 CAAATATGTATCCGCTCATGAGACAATAACCTCATAAATGCTTCAATAATATTGAAAA 5054
Db 121 ggaagatattgatttcaacatttccggtgctgccttattcccttttttgcgcatttt 180
Cp 5053 GGAAGAGTATGAGTATTCACATTTCCGTGTGCGCTTATTCCTTTTTTGGGCGATT 4964
Db 181 gcttccctgttttctcaccacaaacgctggtgaaagtataaagatgctcaaatcagt 240
Cp 4993 GCCTTCCTGTTTTTGTCTCACCAGAAACGCTGGTGAAGATAAAGATGCTGAAGATCAGT 4934
Db 241 tgggtgcacgagtggtttacatcgaactggatctcaacagcggttaaatccttgaagt 300
Cp 4933 TGGGTGCAGAGTGGGTTACATCGAAGTGGATCTCAACACGCTAAGATGCTTGAGATT 4874
Db 301 ttgcgccgaagaacgcttttccaatgatgacacacttttaaaagtctctgtatgagcgg 360
Cp 4873 TTGCGCCGGAAGAGCGTTTTTCCAATGATGAGCACATTTTAAAGTTCTGCTATGTGGCGCG 4814
Db 361 tattatccgctattgaccccggaagcaactcgctcccgccatcacactattctcaga 420
Cp 4813 TATTATCCCGTATTGACGCGCGGCAAGAGCAACTCGGTGCGCGCATACATATTCACAGA 4754
Db 421 atgacttggttgagtaactcaccagtcacagaaagcatctttacgtagtgccatcagtaa 480
Cp 4753 ATGACTTGGTTGAGTACTCACCGATCACAGAAAAGCATCTTACGGATGGCATGACAGTAA 4694
Db 481 ggaattatgagtgctgctccataaccatgagtgataacactgcggcccaacttactctga 540
Cp 4693 GAGAATTATGCACTGCTGCCATAACCATGAGTGAATAACACTTGGCGCAACTTACTTCTGA 4634
Db 541 caacgatcgagagccggaagagctaacccgcttttttgcacaacatggggatcatataa 600
Cp 4633 CAACGATCGGAGGACCGAAGAGAGTAACCGCTTTTTTTCACACACATGSGGATCACTAGTAA 4574
Db 601 ctgccttgatcggttggaacccgagctgaatgaagccatataccaaaacgacgagctgaca 660
Cp 4573 CTCGCTTGATCGTTGGGAACCGGAGCTGAATGAAGCATACCAAAACGAGGAGTGACAA 4514
Db 661 ccacgatgcctgtagcaatggcaacacgcttgccaaactatttaactggggaactactta 720
Cp 4513 CCACGATGCTCTAGCAATGGCAACACAGCTTGCGCAAACTATTAACTGGGGAAGTACTTAA 4454
Db 721 ctctagcttcccgccgaacaataatagactggatggaggggataaaagtgtcagagccac 780
Cp 4453 CTCTAGCTTCCCGGCAACAATTATAGACTGGATGGAGGGGATAAAGTTGCCAGAGCCAC 4394
Db 781 ttctgcgtcggcccttcgcgctggctggtttattgtgtataaatctggagccggtagc 840
Cp 4393 TTCTGCGCTCGGCCCTTCGCGCTGGCTGCTGTTTTATTGTGATAAAATCTGGAGCGGTGAGC 4334
Db 841 gtgggtctcggggtatcattgacagactgggcccagatgataagccctccgctatcag 900
Cp 4333 GTGGGTCTCGCGGTATCATTTGACGACCTGGGGCCAGATGGTAAAGCCCTCCCGGTATC 4274
Db 901 ttatctcacacgaggggagtcaggcaactatggatgaacgaaatagacagactgctaga 960
Cp 4273 TTATCTACACGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTCAGA 4214
Db 961 taggtgcctcactgattaaagcattggttaactgctcagacccaagttaactatatacttt 1020
Cp 4213 TAGGTGCTCTCAGCTGATTAAAGCATTTGTAACGTGTCAGACCAAGTTTACTCATATAGTT 4154
Db 1021 agattgattaaaaacttcattttttaataaaagatctaggtaaatcctttttgata 1080
Cp 4153 AGATTGATTAAAAACTTCTATTITTAATTAAAGAGATCTAGCTGAAGATCCTTTTATATA 4034
Db 1081 atctcatgacaaaaatcccttaacgtgagtttccgttccactgagcgtcagacccccctag 1140
Cp 4093 ATCTCATGACCAAAATCCCTTAAAGTGAGTTTCGTTCCACTGAGTGCACAGCCCGTAG 4034
Db 1141 aaagatcaaaagatctctcagatcccttttttctgcgcgtaaatctgctgcttga 1200

```





CP 4014 CCAAGGGGGAATTAAAGGCTGACTAAGGGAACAAAGGTTGG 2072

Search completed: Tue Mar 16 12:20:37 1999  
Job time : 915 secs.

\*\*\*\*\*  
 WAREHOUSE  
 \*\*\*\*\*  
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.

Copyright (c) 1993-1998 University of Edinburgh, U.K.

Distribution rights by Oxford Molecular Ltd

MPSrch\_nu n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Mar 16 09:49:15 1999; Maspar time 6574.27 Seconds

Tabular output not generated. 1409.082 Million cell updates/sec

Title: >US-09-020-716-7

Description: (1-5173) from US09020716.seq

Perfect Score: 5173

N.A. Sequence: 1 CTAATTGTAAGCTTAATA...

Comp: GATTAAACATTCGCAATTAT...

Scoring table: TABLE default

Gap 6

Mismatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: emb1-est56

Database: genbank-est109

5:gb-est14 6:gb-est10 7:gb-est11 8:gb-est12 9:gb-est13

10:gb-est14 11:gb-est15 12:gb-est16 13:gb-est17

14:gb-est18 15:gb-est19 16:gb-est20 17:gb-est21

18:gb-est21 19:gb-est3 20:gb-est4 21:gb-est5 22:gb-est6

23:gb-est7 24:gb-est8 25:gb-est9 26:gb-est1 27:gb-est2

28:gb-est3 29:gb-est4

Statistics: Mean 12.749; Variance 2.640; scale 4.830

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description	Pred. No.
1	564	12.8	711	26	Homo sapiens genomic D	0.00e+00
2	564	12.8	711	27	Homo sapiens genomic D	0.00e+00
3	557	12.7	721	27	Homo sapiens genomic D	0.00e+00
4	557	12.7	721	26	Homo sapiens genomic D	0.00e+00
5	635	12.3	718	26	Homo sapiens genomic D	0.00e+00
6	535	12.3	718	27	Homo sapiens genomic D	0.00e+00
7	530	12.2	698	26	Homo sapiens genomic D	0.00e+00
8	530	12.2	698	27	Homo sapiens genomic D	0.00e+00
9	624	12.1	696	27	Homo sapiens genomic D	0.00e+00
10	624	12.1	696	26	Homo sapiens genomic D	0.00e+00
11	624	12.1	720	28	Homo sapiens genomic D	0.00e+00
12	524	12.1	720	26	Homo sapiens genomic D	0.00e+00
13	619	12.0	747	26	Homo sapiens genomic D	0.00e+00

C	14	609	11.8	644	28	AQ115013	CIT-HSP-2374C24.TF CIT	0.00e+00
C	15	603	11.7	692	27	AG0040352	CIT-HSP-2327K21.TF CIT	0.00e+00
C	16	603	11.7	695	26	AG002183	Homo sapiens genomic D	0.00e+00
C	17	598	11.6	748	17	A1124281	1.53 Drosophila 8-12 h	0.00e+00
C	18	593	11.5	640	27	AG074298	CIT-HSP-2381L21.TF CIT	0.00e+00
C	19	595	11.5	654	26	B78704	CIT-HSP-A-890F5.TF C	0.00e+00
C	20	594	11.5	690	27	AG009464	Homo sapiens genomic D	0.00e+00
C	21	597	11.5	703	26	AG001761	Homo sapiens genomic D	0.00e+00
C	22	588	11.4	652	28	AQ108633	CIT-HSP-2379M3.TF CIT	0.00e+00
C	23	583	11.3	594	26	C17B6	Clona intestinalis gen	0.00e+00
C	24	583	11.3	566	27	AG079096	CIT-HSP-2356E24.TF CIT	0.00e+00
C	25	577	11.2	689	28	AG014394	Homo sapiens genomic D	0.00e+00
C	26	566	10.9	642	28	AQ113213	CIT-HSP-2372L22.TF CIT	0.00e+00
C	27	566	10.9	682	27	AG074693	CIT-HSP-2301L23.TF CIT	0.00e+00
C	28	557	10.8	615	28	AQ113424	CIT-HSP-2376F2.TF CIT	0.00e+00
C	29	559	10.8	742	28	AQ108392	CIT-HSP-2383E2.TF CIT	0.00e+00
C	30	551	10.7	604	26	AG002706	Homo sapiens genomic D	0.00e+00
C	31	550	10.6	634	27	AQ038010	CIT-HSP-2326F2.TF CIT	0.00e+00
C	32	549	10.5	653	27	AQ007504	CIT-HSP-2365K21.TF CIT	0.00e+00
C	33	544	10.5	633	27	AQ040787	CIT-HSP-2336K22.TF CIT	0.00e+00
C	34	545	10.5	661	27	AQ057603	CIT-HSP-2345M2.TF CIT	0.00e+00
C	35	545	10.5	690	27	AQ029399	RPC111-39F1.TP RPC111	0.00e+00
C	36	536	10.4	626	28	AQ111342	CIT-HSP-2371M2.TF CIT	0.00e+00
C	37	540	10.4	637	27	AQ007360	CIT-HSP-2365E1.TF CIT	0.00e+00
C	38	531	10.3	549	27	AQ003191	RPC111-19E5.TP RPC111	0.00e+00
C	39	534	10.3	594	26	B71605	RPC111-9N1.TP RPC111 H	0.00e+00
C	40	526	10.2	540	27	AQ074307	CIT-HSP-2382M1.TF CIT	0.00e+00
C	41	526	10.2	594	26	AG007826	Homo sapiens genomic D	0.00e+00
C	42	527	10.2	629	27	AG011367	Homo sapiens genomic D	0.00e+00
C	43	530	10.2	630	27	AQ003238	RPC111-19L13.TP RPC11	0.00e+00
C	44	530	10.2	743	26	AG002826	Homo sapiens genomic D	0.00e+00
C	45	525	10.1	600	28	AQ111411	CIT-HSP-2373F2.TF CIT	0.00e+00

#### ALIGNMENTS

RESULT 1 AG004508 711 bp DNA GSS 29-JAN-1998  
 LOCUS Homo sapiens genomic DNA, 21q region, clone: S594BG38, genomic  
 DEFINITION survey sequence.  
 ACCESSION AG004508  
 NID G2822058  
 KEYWORDS GSS.  
 SOURCE Homo sapiens DNA, clone: S594BG38.  
 ORGANISM Homo sapiens  
 Eukaryotae: Metazoa: Chordata: Vertebrata, Mammalia, Eutheria:  
 Primates: Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 711)  
 Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.  
 Published Only in DataBase (1998) In press  
 2 (bases 1 to 711)  
 Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.  
 Direct Submission  
 Submitted (29-JAN-1998) to the DDBJ/EMBL/GenBank databases.  
 Masahira Hattori, Kitasato University, Department of Science, JS-  
 Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan  
 (E-mail: hattori@hgc.ims.u-tokyo.ac.jp, Tel: 0427-78-9732,  
 Fax: 0427-78-9561)

#### FEATURES

Location/Qualifiers  
 1..711  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="21"  
 /clone="S594BG38"  
 /map="21q"

BASE COUNT 169 a 188 c 172 g 182 t

Query Match 12.8%; Score 664; DB 26; Length 711;

Best Local Similarity 99.6%; Pred. No. 0.00e+00;

Matches 68%; Conservative 0; Mismatches 0; Indels 3; Caps 3;





REFERENCE 1 (bases 1 to 721)  
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.  
 TITLE Homo sapiens genomic DNA, chromosome 21q  
 JOURNAL Published Only in Database (1998) In press  
 REFERENCE 2 (bases 1 to 721)  
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (29-JUL-1998) to the DBJ/EMBL/GenBank databases.  
 Masahira Hattori, Kitasato University, Department of Science, JST  
 Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan  
 (E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,  
 Fax:0427-78-9561)  
 FEATURES  
 source Location/Qualifiers  
 1. .721  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="21"  
 /clone="879G12X91"  
 /map="21q"  
 BASE COUNT 195 a 163 c 175 g 179 t 9 others  
 ORIGIN  
 Query Match 12.7% Score 657; DB 27; Length 721;  
 Best Local Similarity 98.7%; Pred. No. 0.00e+00;  
 Matches 683; Conservative 0; Mismatches 5; Indels 4; Gaps 4;  
 Db 32 CATCTTACGGATGGCATGACAGTAAGAGAAATTATGCACTGCTGCCATACCATGAGTAT 91  
 Cp 4718 CATCTTACGGATGGCATGACAGTAAGAGAAATTATGCACTGCTGCCATACCATGAGTAT 4659  
 Db 92 AACACTGCGGCCAATCTTACTTGTGCAAGCATGCGAGGACCGAAGAGCTAACCGCTTTT 151  
 Cp 4658 AACACTGCGGCCAATCTTACTTGTGCAAGCATGCGAGGACCGAAGAGCTAACCGCTTTT 4599  
 Db 152 TTGCACAAATGGGGATCAATGTAATCGCTGTGATCGCTTTGGAAACCGAGCTGAATGAA 211  
 Cp 4598 TTGCACAAATGGGGATCAATGTAATCGCTGTGATCGCTTTGGAAACCGAGCTGAATGAA 4539  
 Db 212 GCGATACCAACAGCAGGAGTGCACACGATGCTGTAGCAATGGCAACACGTTGGCG 271  
 Cp 4538 GCGATACCAACAGCAGGAGTGCACACGATGCTGTAGCAATGGCAACACGTTGGCG 4479  
 Db 272 AACATTTAACTGGCGAATCTTACTTGTGCAAGCATGCGAGGACCGAAGAGCTAACCGCTTTT 331  
 Cp 4478 AACATTTAACTGGCGAATCTTACTTGTGCAAGCATGCGAGGACCGAAGAGCTAACCGCTTTT 4419  
 Db 332 GAGCGGATAAAGTTTCAGGACACCTTCTGCGCTCGGCCCTTCGGCTGGCTGTTATT 391  
 Cp 4418 GAGCGGATAAAGTTTCAGGACACCTTCTGCGCTCGGCCCTTCGGCTGGCTGTTATT 4359  
 Db 392 GCTGATAAATCTGGAGCCGGTGGGTGCTCGCGGTATCATTTGCAGCACTGGGGCCA 451  
 Cp 4358 GCTGATAAATCTGGAGCCGGTGGGTGCTCGCGGTATCATTTGCAGCACTGGGGCCA 4299  
 Db 452 GATGGTAAGCCCTCCGATCTGATGTTATCTACACGACGGGAGTCAGGCAACTATGAT 511  
 Cp 4298 GATGGTAAGCCCTCCGATCTGATGTTATCTACACGACGGGAGTCAGGCAACTATGAT 4239  
 Db 512 GAACGAATATAGACAGATCGGTGAGATAGTGCCTCACTGATTAAAGC-TTGGTAAGTGA 570  
 Cp 4238 GAACGAATATAGACAGATCGGTGAGATAGTGCCTCACTGATTAAAGC-TTGGTAAGTGA 4179  
 Db 571 GACCAAGTTTACTC-TATATACCTTTAGATGATTTAAACCTTCATTTTAAATTTAAAGG 629  
 Cp 4178 GACCAAGTTTACTC-TATATACCTTTAGATGATTTAAACCTTCATTTTAAATTTAAAGG 4119  
 Db 630 ATCTAGGTGAAGATCCTTTT-GATAATCNCATGACCAAA-TCCCCTAAGCTGAGTTTCG 687  
 Cp 4118 ATCTAGGTGAAGATCCTTTT-GATAATCNCATGACCAAA-TCCCCTAAGCTGAGTTTCG 6059  
 Db 688 TTCCACTCAGCGTCAGACCCCGTAGAAAAAT 719  
 Cp 4058 TTCCACTCAGCGTCAGACCCCGTAGAAAAAT 4027

4 AG004662 721 bp DNA GSS 30-JAN-1998  
 Homo sapiens genomic DNA, 21q region, clone: 879G12X91, genomic  
 survey sequence.  
 AG004662  
 NID Q2826191  
 KEYWORDS GSS  
 SOURCE Homo sapiens DNA, clone:879G12X91.  
 ORGANISM Homo sapiens  
 Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
 Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 721)  
 AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.  
 TITLE Homo sapiens genomic DNA, chromosome 21q  
 JOURNAL Published Only in Database (1998) In press  
 REFERENCE 2 (bases 1 to 721)  
 AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (30-JAN-1998) to the DBJ/EMBL/GenBank databases.  
 Masahira Hattori, Kitasato University, Department of Science, JST  
 Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan  
 (E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,  
 Fax:0427-78-9561)  
 FEATURES  
 Location/Qualifiers  
 1. .721  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="21"  
 /clone="879G12X91"  
 /map="21q"  
 BASE COUNT 195 a 163 c 175 g 179 t 9 others  
 ORIGIN  
 Query Match 12.7% Score 657; DB 26; Length 721;  
 Best Local Similarity 98.7%; Pred. No. 0.00e+00;  
 Matches 683; Conservative 0; Mismatches 5; Indels 4; Gaps 4;  
 Db 32 CATCTTACGGATGGCATGACAGTAAGAGAAATTATGCACTGCTGCCATACCATGAGTAT 91  
 Cp 4718 CATCTTACGGATGGCATGACAGTAAGAGAAATTATGCACTGCTGCCATACCATGAGTAT 4659  
 Db 92 AACACTGCGGCCAATCTTACTTGTGCAAGCATGCGAGGACCGAAGAGCTAACCGCTTTT 151  
 Cp 4658 AACACTGCGGCCAATCTTACTTGTGCAAGCATGCGAGGACCGAAGAGCTAACCGCTTTT 4599  
 Db 152 TTGCACAAATGGGGATCAATGTAATCGCTGTGATCGCTTTGGAAACCGAGCTGAATGAA 211  
 Cp 4598 TTGCACAAATGGGGATCAATGTAATCGCTGTGATCGCTTTGGAAACCGAGCTGAATGAA 4539  
 Db 212 GCGATACCAACAGCAGGAGTGCACACGATGCTGTAGCAATGGCAACACGTTGGCG 271  
 Cp 4538 GCGATACCAACAGCAGGAGTGCACACGATGCTGTAGCAATGGCAACACGTTGGCG 4479  
 Db 272 AACATTTAACTGGCGAATCTTACTTGTGCAAGCATGCGAGGACCGAAGAGCTAACCGCTTTT 331  
 Cp 4478 AACATTTAACTGGCGAATCTTACTTGTGCAAGCATGCGAGGACCGAAGAGCTAACCGCTTTT 4419  
 Db 332 GAGCGGATAAAGTTTCAGGACACCTTCTGCGCTCGGCCCTTCGGCTGGCTGTTATT 391  
 Cp 4418 GAGCGGATAAAGTTTCAGGACACCTTCTGCGCTCGGCCCTTCGGCTGGCTGTTATT 4359  
 Db 392 GCTGATAAATCTGGAGCCGGTGGGTGCTCGCGGTATCATTTGCAGCACTGGGGCCA 451  
 Cp 4358 GCTGATAAATCTGGAGCCGGTGGGTGCTCGCGGTATCATTTGCAGCACTGGGGCCA 4299  
 Db 452 GATGGTAAGCCCTCCGATCTGATGTTATCTACACGACGGGAGTCAGGCAACTATGAT 511  
 Cp 4298 GATGGTAAGCCCTCCGATCTGATGTTATCTACACGACGGGAGTCAGGCAACTATGAT 4239  
 Db 512 GAACGAATATAGACAGATCGGTGAGATAGTGCCTCACTGATTAAAGC-TTGGTAAGTGA 570  
 Cp 4238 GAACGAATATAGACAGATCGGTGAGATAGTGCCTCACTGATTAAAGC-TTGGTAAGTGA 4027



Db	141	CATCAAAAATCGACGCTCAAGTCAGAGGTGCGGAAACCCGACAGGACTATAAAGATAC	200
QY	3454	CATCAAAAATCGACGCTCAAGTCAGAGGTGCGGAAACCCGACAGGACTATAAAGATAC	3513
Db	201	CAGGCGTTTCCCCCTGGAAGCTCCCTCGTGGGCTCTCTGTTCCGACCTGCCGTTACC	260
QY	3514	CAGGCGTTTCCCCCTGGAAGCTCCCTCGTGGGCTCTCTGTTCCGACCTGCCGTTACC	3573
Db	261	GGATAGCTTCGCGCCTTTCTCCCTTCGGGAAGGTGCGCCTTCTCATAGCTCAGCGTGT	320
QY	3574	GGATAGCTTCGCGCCTTTCTCCCTTCGGGAAGGTGCGCCTTCTCATAGCTCAGCGTGT	3633
Db	321	AGGTATCTCAGTTCGGTGTAGGTTCGCTCGAAGCTGGGCTGTGTGCAGCAACCCGCC	380
QY	3634	AGGTATCTCAGTTCGGTGTAGGTTCGCTCGAAGCTGGGCTGTGTGCAGCAACCCGCC	3693
Db	381	GTTACGCGGACGCTGGCGCCTTATCGGTAACATATCGCTTTGAGTCCAAACCCGTAAGA	440
QY	3694	GTTACGCGGACGCTGGCGCCTTATCGGTAACATATCGCTTTGAGTCCAAACCCGTAAGA	3753
Db	441	CAGGACTTATGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGACNAGGTATGTA	500
QY	3754	CAGGACTTATGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGACGAGGTATGTA	3813
Db	501	GGCGGTGCTAACAGAGTCTTTGAAGTGGTGGCCTAACTAGCGCTAACACTAGAAGACAG	560
QY	3814	GGCGGTGCTA-CAGAGTCTTTGAAGTGGTGGCCTAACTAGCGCTA-CACTAGAAGGACAG	3871
Db	561	TATTTGGTATCTGCCCTCTGCTGAAGCCAGTTANCTTTCGGAAAGAGTTGGTAGCTCT	620
QY	3872	TATTTGGTATCTGCCCTCTGCTGAAGCCAGTTACCTT-CGGAAAAAGAGTTGGTAGCTCT	3930
Db	621	TGATCGGGGAAAACAACCAAGCTGGTAGCG-1GTTTTTTTGTTTTGCAGCAGCAGATA	679
QY	3931	TGATCGGGGAAAACAACCAAGCTGGTAGCGTGGTTTTTTTGTTCGCAAGCAGCAGATT	3990
Db	680	ACGGCAGAAAAAAGGATCTCAAGA-GATC-TTTGATCT 718	
QY	3991	ACGGCAGAAAAAAGGATCTCAAGAAGATCTCAAGAAGATCTTTGATCT 4031	

RESULT	7			
LOCUS	AG003787	698 bp	DNA	GSS
DEFINITION	Homo sapiens genomic DNA, 21q region, clone: T485XN, genomic survey sequence.			
ACCESSION	AG003787			
NID	g2734689			
KEYWORDS	GSS.			
SOURCE	Homo sapiens DNA, clone:T485XN.			

SOURCE: Homo sapiens DNA, Clone:14653XN.  
 ORGANISM: Homo sapiens  
 Eukaryotae; Metazoa; Chordata; Vertebrata; Eutheria;  
 Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 698)  
 Hattori.M., Ishii.K., Shiba.T. and Sakaki.Y.  
 TITLE: Homo sapiens genomic DNA, chromosome 21q  
 JOURNAL: Published Only in DataBase (1998) In press  
 2 (bases 1 to 698)  
 Hattori.M., Ishii.K., Shiba.T. and Sakaki.Y.  
 TITLE: Direct Submission  
 JOURNAL: Submitted (07-JAN-1998) to the DDBJ/EMBL/GenBank databases.  
 Masahira Hattori, Kitasato University, Department of Science, JST  
 Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan  
 (E-mail:hattori@hgc.ims.u-tokyo.ac.jp; tel:0427-78-732,

FEATURES	source	Location/Qualifiers
tax:0427069501)		
	1. .698	
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/chromosome="21"
		/clone="T485XN"
		/map="21q"
PASE COUNT	189 a	133 c 171 g 175 t 10 others
ORIGIN		

Query Match	12.2%	Score 630;	DB 26;	Length 698;
Best Local Similarity	99.0%;	Pred. N. 0.00e+00;		
Matches	661;	Conservative	0;	Mismatches 2; Indels 5; Gaps 5;
Db	31	CTTAGCGTGGCATGACAGTAAGAGTAATTATGAGTGTCTGCCATAACCATTAGTGTATAC	90	
Cp	4715	CTTACGGATGGCATGACAGTAAGAGTAATTATGAGTGTCTGCCATAACCATTAGTGTATAC	4655	
Db	91	ACTGGGGCCAACTTACTTCTGACAACGATGAGAGCGCGAAGGAGTAACCGCTTTTITG	150	
Cp	4655	ACTGGGGCCAACTTACTTCTGACAACGATGAGAGCGCGAAGGAGTAACCGCTTTTITG	4596	
Db	151	CACACATGGGGATCATGTAACTCGGCTTGTATCGTTGGACCGGAGCTGAATGAAGCC	210	
Cp	4595	CACAACATGGGGATCATGTAACTCGGCTTGTATCGTTGGACCGGAGCTGAATGAAGCC	4536	
Db	211	ATACCAACGAGGAGCGGTGACACACCATGCTGTAGCAATGCAACAACGTTGGCGAAA	270	
Cp	4535	ATACCAACGAGGAGCGGTGACACACCATGCTGTAGCAATGCAACAACGTTGGCGAAA	4476	
Db	271	CTATTAACCTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAG	330	
Cp	4475	CTATTAACCTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAG	4415	
Db	331	GGGATAAAGTTGCAGGACCACTTCTCGGCTCGGCCCTTCCGGCTGGCTGTTTATTCGT	390	
Cp	4415	GGGATAAAGTTGCAGGACCACTTCTCGGCTCGGCCCTTCCGGCTGGCTGTTTATTCGT	4356	
Db	391	GATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTCACGACTGGGGCCAGAT	450	
Cp	4355	GATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTCACGACTGGGGCCAGAT	4296	
Db	451	GGTAAGCCCTCCCGTATCGTATAGTTATCTACACGACGGG-AGTCAGGCAACTATGGATGAA	509	
Cp	4295	GGTAAGCCCTCCCGTATCGTATAGTTATCTACACGACGGGAGTCAGGCAACTATGGATGAA	4236	
Db	510	CGAATATAGACATTCGCTGAGATAGTGGCTCACTCATTAAGCAATGGTAACTGTCAGA	569	
Cp	4235	CGAATATAGACATTCGCTGAGATAGTGGCTCACTCATTAAGCAATGGTAACTGTCAGA	4177	
Db	570	CCAAGTTTACTC-TATATACTTTAGATTGATTTAAAACTTCATTTTAAATTTAAAAGAT	628	
Cp	4176	CCAAGTTTACTCATATATACATTAGATTGATTTAAAACCTTCATTTTAAATTTAAAAGAT	4117	
Db	629	CTAGGTGAAGATCNTTTT-GATAATCTCATGACCAAAATCCCTTACGCTGAGTTTCGTT	687	
Cp	4116	CTAGGTGAAGATCCTTTTGTATAATCTCATGACCAAAATCCCTTACGCTGAGTTTCGTT	4057	
Db	688	TCCACTGA 695		
Cp	4056	-CCACTGA 4050		
RESULT	8	AG009976	698 bp	DNA
LOCUS		Homo sapiens	genomic DNA, 21q	region, clone: T485XN, genomic survey
DEFINITION			sequence.	
ACCESSION		AG009976		
NID		93289962		
KEYWORDS		GSS.		
SOURCE		Homo sapiens		
ORGANISM		Homo sapiens		
REFERENCE		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;		
AUTHORS		Primates; Catarrhini; Hominiidae; Homo.		
TITLE		1 (bases 1 to 698)		
JOURNAL		Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.		
REFERENCE		2 (bases 1 to 698)		
AUTHORS		Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.		
TITLE		Direct Submission		
JOURNAL		Submitted (06-JUL-1998) to the DDBJ/EMBL/GenBank databases.		



Cp 4388 CGCTGGCCCTT-CGGCTGG-CTGGTTTATGCTGATAAATCTGGAGCCCGGTGAGCGTG 4331

Db 690 GGTCTC 695  
|||||

Cp 4330 GGTCTC 4325

RESULT 10

LOCUS	AG003576	696 bp	DNA	GSS	19-DEC-1997
DEFINITION	Homo sapiens genomic DNA, 21q region, clone: P8G4SpN15, genomic survey sequence.				
ACCESSION	AG003576				
NID	q2706702				
KEYWORDS	GSS.				
SOURCE	Homo sapiens DNA, clone:P8G4SpN15.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 696)				
TITLE	Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.				
JOURNAL	Homo sapiens genomic DNA, chromosome 21q				
REFERENCE	Published Only in DataBase (1997) In press				
AUTHORS	2 (bases 1 to 696)				
TITLE	Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.				
JOURNAL	Direct Submission				
REFERENCE	Submitted (19-DEC-1997) to the DDBJ/EMBL/GenBank databases.				
AUTHORS	Masahiro Hattori, Kitasato University, Department of Science, Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan				
TITLE	(E-mail: hattori@pc.ims.u-tokyo.ac.jp. Tel:0427-78-9732, Fax:0427-78-9561)				
FEATURES	Location/Qualifiers				
source	1..696				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/chromosome="21"				
	/clone="P8G4SpN15"				
	/map="21q"				
BASE COUNT	183 a	165 c	185 g	162 t	1 others
ORIGIN					
Query Match	12.1% Score 624, DB 26; Length 696;				
Best Local Similarity	99.1% Pred No. 0 00e+00;				
Matches	650:	Conservative	0:	Mismatches	0: Indels 6: Gaps 6:
Db	31	GTTTTGTCTACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGCA	90		
Cp	4985	GTTTTGTCTACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGCA	4926		
Db	91	CGAGTGGGTTACATCGAATCGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCCCC	150		
Cp	4925	CGAGTGGGTTACATCGAATCGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCCCC	4865		
Db	151	GAAGAACGTTTTCGAATGATGAGCACTTTTAAAGTTTCTGCTATGTGGCGCGGATTATCC	210		
Cp	4865	GAAGAACGTTTTCGAATGATGAGCACTTTTAAAGTTTCTGCTATGTGGCGCGGATTATCC	4806		
Db	211	CGATTATGCGCGGGGAGAGCACTCGGTGCGCGGATACACTATTCTCAGATGACTTG	270		
Cp	4805	CGATTATGCGCGGGGAGAGCACTCGGTGCGCGGATACACTATTCTCAGATGACTTG	4746		
Db	271	GTTGAGTACTCACAGTACAGAAAGCACTTTACGGATGGCATCACAGTAAAGAGAATTA	330		
Cp	4745	GTTGAGTACTCACAGTACAGAAAGCACTTTACGGATGGCATCACAGTAAAGAGAATTA	4586		
Db	331	TGCAGTGTGCCATAACCATGAGTGATAACACTTGGCGGCACTTACTTCTGACAACGATC	390		
Cp	4685	TGCAGTGTGCCATAACCATGAGTGATAACACTTGGCGGCACTTACTTCTGACAACGATC	4626		
Db	391	GGAGGACCGAAGAGCTAACCGCTTTTTTGCACAAGATGGGGATCATGTAACTCGCCTT	450		
Cp	4625	GGAGGACCGAAGAGCTAACCGCTTTTTTGCACAAGATGGGGATCATGTAACTCGCCTT	4566		
Db	451	GATCGTTGGGAACCGAGCTGAATSAAGCCATAACCAACGACGAGCGTGAACACACGAT	510		



```

/chromosome="21"
/clone="c48B2A23"
/map="21q"

BASE COUNT 194 a 171 c 176 g 194 t 12 others
ORIGIN

Query Match 12.0%; Score 619, DB 26, Length 747;
Best Local Similarity 96.3%; Pred. No. 0.00e+00;
Matches 683; Conservative 0; Mismatches 17; Indels 9; Gaps 9;

Db 41 ATCTGGGCTCTGCTGAAGCAGTATCTTCGAAAAAGAGTGTAGCTCTTGATCGGC 100
QY ATCTGGGCTCTGCTGAAGCAGTATCTTCGAAAAAGAGTGTAGCTCTTGATCGGC 3939
Db 101 AAACAAACACCGCTCGTAGCGGTG-TTTTTTTTGTTCAGCAGCAGATACCGGAGA 159
QY AAACAAACACCGCTCGTAGCGGTGTTTTTTTGTTCAGCAGCAGATACCGGAGA 3999
Db 160 AAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAAC 219
QY AAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAAC 4059
Db 220 GAAACATCAGTTAAGGATTTTGTTCATGAGATTATCAAAAGAGATCTTCACTAGATC 279
QY GAAACATCAGTTAAGGATTTTGTTCATGAGATTATCAAAAGAGATCTTCACTAGATC 4119
Db 280 CTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATCAGTAAACTTGGTCT 339
QY CTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCT 4179
Db 340 GACAGTTACCAATGCTTAATCAGTAGAGGACCTATCTCAGCGATCTGCTATTTCGTCA 399
QY GACAGTTACCAATGCTTAATCAGTAGAGGACCTATCTCAGCGATCTGCTATTTCGTCA 4239
Db 400 TCCATAGTTCCTGACTCCCGTCGTTAGTAGAATCACTAGATACGGAGGCT-ACCATCT 458
QY TCCATAGTTCCTGACTCCCGTCGTTAGTAGAATCACTAGATACGGAGGCTTACCATCT 4299
Db 459 GGCCCGCAGTCTCAATGATACGGGAGAGCCAGCTACCGGCTCCAGATTTATCAGCA 518
QY GGCCCGCAGTCTCAATGATACGGGAGAGCCAGCTACCGGCTCCAGATTTATCAGCA 4359
Db 519 ATAAACAGCAGCCGAGAGGCGGAGGCA-AGTGTCTGCTGAACATTTATCGGCTC 577
QY ATAAACAGCAGCCGAGAGGCGGAGGCGGAGGCA-AGTGTCTGCTGAACATTTATCGGCTC 4418
Db 578 CATCCAGTCTATTAAATGTTGCGGGAAGCTAGAGTAAAGTAGTTTCCGCAAGTTAATAGT 637
QY CATCCAGTCTATTAAATGTTGCGGGAAGCTAGAGTAA-CTAGTT-CCGCAAGTTAATAGT 4476
Db 638 TTGGCAACGTTGTTGNCATGCTGAGGGATCAGTGTGTCAGGCTCCGCTTGGGTAA 697
QY TTGGCAACGTTGTTGNCATGCTGAGGATCAGTGTGTCAGGCTCCGCTTGGGTAA-T 4535
Db 698 GGCTTCATTTCAAGTCCGGTCCCAACGATCAAGGGGTTAAATGAT 746
QY GGCTTCATTT-CAAGTCCGGG-TTCCCAACGATCAAGGGGAGTTACATGAT- 4582

RESULT 14
LOCUS A0115013 644 bp DNA GSS 28-AUG-1998
DEFINITION CIT-HSP-2374C24.FR CIT-HSP Homo sapiens genomic clone 2374C24,
genomic survey sequence.
ACCESSION A0115013
NID g3491134
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 644)
AUTHORS Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,
Berry, K., Granger, D., Suh, E., Wibic, C., Shizuya, H., Simon, M. and

```

```

Venter, J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
Unpublished (1998)

Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel.: 301 838 0200
Fax: 301 838 0208
Email: mdadams@igr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
Location/Qualifiers
source 1..644
/organism="Homo sapiens"
/note="Vector: pBelBAC11; Site_1. HindIII; Site_2:
HindIII"
/db_xref="taxon.9606"
/clone="2374C24"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
BASE COUNT 135 a 183 c 145 g 179 t 2 others
ORIGIN

Query Match 11.8%; Score 609, DB 28, Length 644;
Best Local Similarity 99.8%; Pred. No. 0.00e+00;
Matches 615; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 5 TCACCTGGCGCTCGTTTACACAGCTGCTGACTGGGAAACCCGCTTACCAACCTTAAT 64
Cp 618 TCACCTGGCGCTCGTTTACACAGCTGCTGACTGGGAAACCCGCTTACCAACCTTAAT 559
Db 65 CGCCTTGACGACATCCCGCTTTCCGCGAGCTGGGTAAATAGGAAGAGGCGGACCGAT 124
Cp 558 CGCCTTGACGACATCCCGCTTTCCGCGAGCTGGGTAAATAGGAAGAGGCGGACCGAT 499
Db 125 CGCCTTCCCAACAGTTGGCAGCCTGAATGGGAATGG-ACGGCGCCTGTAGCGGGCA 183
Cp 498 CGCCTTCCCAACAGTTGGCAGCCTGAATGGGAATGGGACGCGCTGTAGCGGGCA 439
Db 184 TTAAGCGCGCGGGTGTGGTGTACGGCAGCGGTGACCGGTACACTTCCAGCGGCCA 243
Cp 438 TTAAGCGCGCGGGTGTGGTGTACGGCAGCGGTGACCGGTACACTTCCAGCGGCCA 379
Db 244 GGCCCGCTCCCTTTCCGCTTTCTCCCTTCCCTTCTCGGCAAGTTGGGCGGTTCGCGGT 303
Cp 378 GGCCCGCTCCCTTTCCGCTTTCTCCCTTCCCTTCTCGGCAAGTTGGGCGGTTCGCGGT 319
Db 304 CAGCTCTTAATCGGGGCTCCCTTTTASGGTTCGGATTAGTGTTTACGACCTGAC 363
Cp 318 CAGCTCTTAATCGGGGCTCCCTTTTASGGTTCGGATTAGTGTTTTACGACCTGAC 259
Db 364 GCGAAAAACTTGAATAGGTCATGTTAGGTCAGTGTAGGCGCATGCGCTGATACAGGTT 423
Cp 258 GCGAAAAACTTGAATAGGTCATGTTAGGTCAGTGTAGGCGCATGCGCTGATACAGGTT 199
Db 424 TTTCCGCTTTTACGCTTGGAGTCCACGTTCTTTTATAGTGACCTCTTGTTCACAACTGA 483
Cp 198 TTTCCGCTTTTACGCTTGGAGTCCACGTTCTTTTATAGTGACCTCTTGTTCACAACTGA 139
Db 484 ACAACACTCAACCTATCTCGGTCTATTCTTTTGTATTTATAGGAGTTTCCCGATTCG 543
Cp 138 ACAACACTCAACCTATCTCGGTCTATTCTTTTGTATTTTATAGGAGTTTCCCGATTCG 79
Db 544 GCCTATTGTTAAAAAATCAGCTGATTTTAAACAAAATTTTAAACGGAATTTTAAACAAATA 603
Cp 78 GCCTATTGTTAAAAAATCAGCTGATTTTAAACAAAATTTTAAACGGAATTTTAAACAAATA 19

```







1

1

---





VCQAIIIDQRREIRIELMQLKNEIIVCMELYSALQRLKLERDFTWDEYIEYLKSVNPQ  
 LSSQAEVMEHSDYVVRVKNLEQVAFMALVIMVDFADERSDCVFKLAKFKGV  
 LSSQAEVMEHSDYVVRVKNLEQVAFMALVIMVDFADERSDCVFKLAKFKGV  
 TLPHRYREGHEMETRATAVOYANDIAHSEHDLVLRGAVGSKSLGPHVLSVAGSV  
 QLAIEFRFDECHVLFQSSSEFPKPLRMFGNSIFGSSPISVMTSGFALHYFANPNS  
 LIQVHQAAATSLAKDLKLGKWNKSLVAKOLIIAGAVAGIGLISWFTQSVETVSH  
 QGKNKXPIQALKPHAPDKRAGEIDNDDTIEFFGSAVRKKGKGGKTVGMGKSH  
 RRFIMNGFDPTESFIOFVDPDLTGRIEENYADIDIOERSEVRKMKVENDDIEM  
 GRMKHGLEIVTIGTSGARPHFVATNIENGVTLDIDVVDVFLGVSPFLDIDNRSI  
 YNKVSYSGRIORLGRVGRKGVALLRGTEKGIIEIPSMVATAALACAFYNLP  
 VMTGSVSLIGCTVQVMTQOFLSPFFIONFVAHDSMPVITHDLKYLKLDRC  
 MTPCDQSPRASSTWLSVSEYEPGLVALEIPQVIAFHIEIPKLEHMLWETV  
 KYKCVLPFSIRASSISKIAITLTDLFAIPRTLILVLEERVKQSEFRSLIDEG  
 CSMFVNLNTLRARYAKDYTAENIKLEKRSQLEFNSLDGSACEENLKRYES  
 LQVHQAAATSLAKDLKLGKWNKSLVAKOLIIAGAVAGIGLISWFTQSVETVSH  
 QGKNKXPIQALKPHAPDKRAGEIDNDDTIEFFGSAVRKKGKGGKTVGMGKSH  
 RRFIMNGFDPTESFIOFVDPDLTGRIEENYADIDIOERSEVRKMKVENDDIEM  
 QALGNTIHAIFKDCDKALKIDLPHPLKVCCKTNGIAFPFERELEROTGPV  
 EVDVDPQAEVEHAKSLMGLRDFNPIDQVLCRLKSVSEYGAEMVGFQAYIVA  
 ADKEIIVMSCLRLYGLLGIWNSLAEURKEKILANKIRTFITAAPDLTLGKVC  
 VDDNNQYSKIECCVTGNTKFGYGMOKLLRRLPENWYVYDADGSOFSLSPTVLI  
 NAVLTIRSTYMEDVDGLQNLNLYTEIVTPISTPDGTIVKFRGNSSOPSTVVDN  
 SLVVLMHAYALIKCEVEFEIDSTCVFVNGDOLLIIVAPKEKESIIDRMSQHFSDLG  
 LNYDFSSTRKEBELWMSHRLGIEGMVYPKLEEBRIVSLOWDRADLEHRLAIC  
 AAMIESGYSLETHQIRRYFSLWLOQOFPATIAQEGKAPYIASMALRKLMDRAVDEE  
 ELRAFTENMVADEFELDSYEVHQAANDIDAGSNNKDKAPQEQSIQPNKKGDK  
 DVNAGTHTVPRKAIKTSKMRPTSKGATVPNLEHLEYAPOQIDISNTRATQSF  
 DTWEAVRMAYDIDETEMPTVMGLMVMWCENGTSPPNVVVMVMDNEOVEXPLAKPI  
 VENAKPTQLRIMAFHSDVAEAYIDMRNKPEMPRYGLIRLNRDMGLARYAFDFEVT  
 SRTPVAREAHIQMAKALKSAQPRFLCLOGISTQENTERTHTEDVSFMSHTLGV  
 KNN.

mat\_peptide 185..1009  
 old\_sequence /note="putative extreme 5'protein"  
 788  
 /citation="[1]  
 /replace=""  
 old\_sequence 847  
 /citation="[1]  
 /replace="c"  
 1010..2656  
 /note="putative"  
 /product="helper component protein, HC"  
 2657..3655  
 /note="putative"  
 /product="38k protein"  
 3656..5557  
 /product="cytoplasmic inclusion protein"  
 5558..5713  
 /note="putative"  
 /product="genome linked protein, Vpg"  
 5714..7009  
 /note="putative"  
 /product="nuclear inclusion protein, Nla"  
 7010..8572  
 /note="putative"  
 /product="nuclear inclusion protein, Nib"  
 8573..9373  
 /note="putative"  
 /product="capsid protein"  
 3004 a 1818 c 2273 g 2609 t

BASE COUNT 3004 a 1818 c 2273 g 2609 t  
 ORIGIN

Query Match 93.2%; Score 796; DB 34; Length 9704;  
 Best Local Similarity 96.8%; Pred. 0.00e+00;  
 Matches 825; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Db 617 AACCTTTGAAGATTGATTGATGCCACATAACCCACTCAAACTTTGTCGACAAAACAATG 6176

QY 1 AACCTTTGAAGATTGATTGATGCCACATAACCCACTCAAACTTTGTCGACAAAACAATG 60  
 Db 6177 GCATTGCCAAATTTCTTGAGAGAGAGCTGGAAGTGAAGAGAGAGCTGGGCGAGCTGTAGAAG 6236  
 QY 61 GCATTGCCAAATTTCTTGAGAGAGAGCTTGAAGTGAAGAGAGAGCTGGGCGAGCTGTAGAAG 120  
 Db 6237 TCGATGTGAAGAGACATACACGAGGAGGTGAGGATGAAGCTAAATCGCTCATGAGAG 6296  
 QY 121 TCGAGTGAAGAGACATACACGAGGAGGTGAGGATGAAGCTAAATCGCTCATGAGAG 180  
 Db 6297 GCTTGAGAGACTTCAACCCAAATTTGCCAAACAGTTTGAAGCTGAAGATATCTGTTGAAT 6356  
 QY 181 GCTTGAGAGACTTCAACCCAAATTTGCCAAACAGTTTGAAGCTGAAGATATCTGTTGAAT 240  
 Db 6357 ATGGGACATCAGAGATGTACGGTTTGGATTTTGAGACATACATAGTAGGAACCCACCAT 6416  
 QY 241 ATGGGACATCAGAGATGTACGGTTTGGATTTTGAGACATACATAGTAGGAACCCACCAT 300  
 Db 6417 TATTAGGAGTTTCAATGGTTCATGGAGGTGCAATGCCATGCCAGGTACATTCAGGGTGA 6476  
 QY 301 TGTTCAGGAGTTTCAATGGTTCATGGAGGTGCAATGCCAGGTACATTCAGGGTGA 360  
 Db 6477 AGAATCTACAGTTTGAAGGTTTCTGCCAAATTAAGGTAGGGACATCATCTCATCAAAA 6536  
 QY 361 AGAATCTACAGTTTGAAGGTTTCTGCCAAATTAAGGTAGGGACATCATCTCATCAAAA 420  
 Db 6537 TGGCGAAGATTTCCCTGCTTTCCACAGAAATTTGCAATTTCCAGCTCTTACACAGAG 480  
 QY 421 TGGCGAAGATTTCCCTGCTTTCCACAGAAATTTGCAATTTCCAGCTCTTACACAGAG 480  
 Db 6597 AAAGAATTTCTTAGTTGGAACCAACTTCCAGAGAGAGTATGCTTCGTGATCATCACAG 6656  
 QY 481 AAAGAATTTCTTAGTTGGAACCAACTTCCAGAGAGAGTATGCTTCGTGATCATCACAG 540  
 Db 6657 AAACAAGCATTCTACATATACAGGAGGACACATCTTGAAGAGCATTGGATTGAACAG 6716  
 QY 541 AAACAAGCATTCTACATATACAGGAGGACACATCTTGAAGAGCATTGGATTGAACAG 600  
 Db 6717 ATATGGACATTGTGGACTACGAGTGTGAGACACCGGATGATGATAGTCGGAATTC 6776  
 QY 601 ATATGGACATTGTGGACTACGAGTGTGAGACACCGGATGATGATAGTCGGAATTC 660  
 Db 6777 ACAGTCTGGCAAAATATGACACACACCACTACTACTACTAGCTTCGATGAAGATTITG 6836  
 QY 661 ACAGTCTGGCAAAATATGACACACACCACTACTACTACTAGCTTCGATGAAGATTITG 720  
 Db 6837 AAAGCAAGTACCTCCGACCAATGAGCACAATGAATGGGTCAAGCTCTGGGTTTAAATC 6896  
 QY 721 AAAGCAAGTACCTCCGACCAATGAGCACAATGAATGGGTCAAGCTCTGGGTTTAAATC 780  
 Db 6897 CAGACAGTGTGTGGGGGGCGCTTGAAGCTTAAAGACAGCAGCTCCGAAAGGTTATCA 6956  
 QY 781 CAGACAGTGTGTGGGGGGCGCTTGAAGCTTAAAGACAGCAGCTCCGAAAGGTTATTA 840  
 Db 6957 AAACAACAAGCTT 6970  
 QY 841 AAACAACAAGCTT 854

RESULT 4  
 LOCUS A08776 9705 bp RNA PAT 27-NOV-1993  
 DEFINITION Potato virus Y complete genome RNA.  
 ACCESSION A08776  
 NID 9492885  
 KEYWORDS  
 SOURCE Potato virus Y.  
 ORGANISM Potato virus Y.  
 Viruses; serNA positive-strand viruses, no DNA stage; Potyviridae;  
 Potyvirus.  
 REFERENCE 1 (bases 1 to 9705)  
 AUTHORS  
 TITLE COMPLETE NUCLEOTIDIC SEQUENCE OF THE COMPLEMENTARY DNA OF THE  
 GENOMIC RNA OF THE POTYVIRUS, GENES CODING FOR THE POTYVIRUS

ORGANISM	REFERENCE
Potato virus Y	1 (bases 1 to 9698)
Viruses, ssRNA positive strand viruses, no RNA stage, potyviral polyvirus	Singh, M. and Singh, R. P.
	Nucleotide sequence and genome organization of a Canadian isolate of the common strain of potato virus Y (PVYc)

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REMARK  
COMMENT  
FEATURES  
source

Can. J. Plant Pathol. 18, 209-214 (1996)  
2 (bases 1 to 9698)  
Singh, R.P.  
Direct Submission  
Submitted (10-MAY-1994) Rudra P. Singh, Fredericton Research  
Centre, Agriculture and Agri-Food Canada, 850 Lincoln Road,  
Fredericton, New Brunswick, E3B 4Z7, Canada  
3 (bases 1 to 9698)  
Singh, R.P.  
Direct Submission  
Submitted (25-MAR-1997) Rudra P. Singh, Fredericton Research  
Centre, Agriculture and Agri-Food Canada, 850 Lincoln Road,  
Fredericton, New Brunswick, E3B 4Z7, Canada  
Sequence update by submitter  
On Mar 25, 1997, this sequence version replaced gi:602403.

Location/Qualifiers  
1..9698  
/organism="Potato virus Y"  
/strain="common strain"  
/specific\_host="Solanum tuberosum L."  
/db\_xref="taxon:12216"  
/lab\_host="Nicotiana tabacum L."  
/clone="P07"  
/tissue-type="leaf"  
125..9370  
/note="nuclear inclusion protein Nib and capsid protein"  
/codon\_start=1  
/product="polyprotein"  
/db\_xref="PDB:1G90388"  
/translation="METISTQGAISSLPNTLKSSMATYMTICGFSPECKLPYSPAC  
GLIVKREVLAVNPFADELTQLSARLLKQYAYRVLKNGTETTYRYKVTDAQIKRIQK  
KLKEDQEFHMAAPASIVSKITTAGGDPKSEPOAQRGILHTPRVRYKMTIRIZ  
KLTEGQMHFIKQVIMSEKRSVHLINKTTHVQYKEILGAYSAIVRAHMMGR  
RVDQDMVTGLLQRLARTKSNQVRIINRRGDSVILNTKSLKGFHSGDGLF  
IVRSGHEGLDAPRSVTOVSDLSMIOFNSADNFWKGLDGNWARMYPSDHTCVAGLP  
VDCGVAALMAHILPCYKICTCAQOAYSLPVSLFKLLKHARDGNLRGAKND  
RFLVKNFLIAHLETPVDLNLLELFELFKSIGKQOAPKLNVLNNFLKCKNT  
AHQVAQSLLELAFQKRTDNKKGDISFRNKLAFKANWNLVLSQNDKKNAN  
FLMGQREYHAKFRSFFEEIDPAKGYSAEYLRKPSGTRKLSIGNVLVDLAEPRQ  
KMGQDYKQPGVRSKTSKNGYVPCCTDLDGSAIESITFYPTPKKHLVTONSGD  
QKFDVPLCGDSEMLYIAQGYCYINFLAMHNTGEEADKDFTKKQDMCVKLGATWP  
TMDLATACTGMDLFPYDVAELPSLVDHDTQCHVDVSFGSTGYHLKASSVS  
QLILFANDELESDIKHYRGVGNACPELPGSTISPRGEGVIMSESAALKLLKGFPR  
PKYMRQLLDEPVLILSLSPCLLMANNNGIFELAVRLWNEKQSLAMIASLAL  
ALRVSAEQLDRAQLIDRAADLDAACCGENLHLYPTALVLYVKNRNECDLFL  
FKAGFPSTYNTVOIMEKNFLSLDDAWKDLWREKLSATWYSYRATESITRYIKPTG  
RADGLKYNISPAQFLGRSAQVYKGTAGSLSERFNINIKYCNISSEFFIRIRFRPL  
TFVTFVNSLLVLSMTSVAVCAQIILDQKRYREIELMQIEKNEIVCMELYSQIRK  
LERDFTWDEYIEVLKSNVQIIVQAOQMEYDVRHSTFCVKNLEOVAFMALVIM  
VEDAERSDCVFKTLNFKGVLSSLDHEVRHQSLDDVIRKFNDENEIDFELSEDITRT  
SSVLDTFESDMORQIOMGHTLPHYRTEGHFMFEFTATAVOVANDIAHSEHLDLVRG  
AVGSGKSTGLPVLHLSVAGSVLLIEPTPLAENVKOLSEPEFKKPTLRMGNSIFGS  
SPISVMTSGEALHYFANNRSQALQNFVIFDSCVLDPSMAERSLSVYHOACKVLK  
VSATPVGREVEFTTQOPVKLIVEDLSFQSDVDAQSKTNADVQFGSNVLYVSSYN  
EVDTLAKLVDKNMVTVDGRTMKHGCLEIVTKGTSARPEHVAWATNIENGVLID  
VWDFGLKVSPLFDINDRSIAYNKESYSGERTQRLRGVRFKGVALLRIGTEKGI  
EIPSWATEALACAFAYLPVMTGGVSTLIGNCTVQVQVTKQOEFELSPFTIQNFH  
DGSWHPVTHLLKAYLKRCQMTPLCQDSIPYRASCWLSVSEYERLVALEIPQIKI  
AFHKEIPPKLHEMLWETVYKVDCLFSPISIRASSISKIATYLRDLFAIPRLILVE  
RLLEERVQSQFSLRSLDEGSSNFVNLNTLRLARYADYTAENIKLEKRSOLK  
EFSNLDGACENILKRYESLQVHHQAATSLAKDLKLGATWKKSLVAKDLIIAGAVA  
IGGILGYSWFTQSVETVSHQGNKSKRIQALFRHARDKRAFEDINDNENYIADIR  
SAYRKKGKGTVMGKSSRKFIMNYFDPTETYSFIQVDPFLTGADEENYIADIR  
IQDRESEVRKKNWENDDIEMQALSNITIHAYFRKDSKALKIDLPNHPNLCVCKDT  
NGIAFERRELELQTPAVENVNVDIPAOEVEHEAKSLMRGLDLPNFAOTVCRKY  
SVEYGTSEWYGFGEAKYIAHHILFRSNGSMVQSHGTFRVKNLSLSVLPKGRD  
ILILKMPDKPVPFQKLFHFRAPIONERVCLVGNFOEKYASSIITETSTYINIPGST  
WKHLETUNGHCGLPVSTADGCLVGHSLNANHTYNSAEDFESKYLIRANEHN  
EWKWSKNYNDPTLVMLGPKLKDSTPKGLFTKTLVDQLEIHDVWPMQAKSHWMEAL  
TGNLQAVATMKSQLVTKRVHGECHFEKFTLVDAEAEAFFRPMVDYQKSLNRLKAA  
IKDINKYSKIDVIGDQDAFEAEINRYLYQVHFKKAYVTDQEFKALNKKAA  
VGMAYGKKDKYFHEFTDAKEEIVMQSLRILYKGLLIGWNSLKAELRCKKILNKA  
TRFTAAPLDILLGKVKVDVDFNQFYSKNIECCVTGMYKEYGWDKLLRLPENWV

YCDADGSOSSLTPLYFNNAVLIIRSTYMEDWDVGLQMLRNLYTEIVVPISTPDGTI  
VKFRGNSQPSSTVNDNSLMVLYAMHYAFIREGIEEFIEEISITVFFVNGDILLIAYN  
IKEDILDRLSQHSFSDLGLNYDFSRTNRKEELWFMSHRLIEGMYVPKLEERIVS  
TLKORADLAERLEAICARMIESWGYSELTHQIRRFYSWLLQOQFPATIDAGEKAPY  
TASNAKLKLYMDRAVDEELRAFETMMVALDDDEFEDSEYVYHQANDIDAGGSNKKD  
TKQSSQISQPNKSGKDKDYNAGTSGTHTVPIKJTSKMRMPKSKGAVALNLEHLE  
YAPQOISNTRAIQSOFDWEYAVPMAYDIGETEMPTVMNGKVCINGENISPNVNG  
VWMMGNEQVEYPLKPIVENAKPTLQINAHESDVAEAYIEMRNKKPEYMPRYGLIR  
NLRDMGLAARYAFDEYVTSRTPPVAREAQIOMRAAALKSNQPLRFLGDLGGISTOENT  
ERHTTEDVSPSMHILLGVKNM"

mat\_peptide 182..1012  
/product="P1"  
1010..2407  
/product="HC"  
2405..2659  
/product="HC-pro"  
2657..3502  
/product="P3"  
3500..3658  
/product="6K1"  
3656..5557  
/product="C1"  
5558..5716  
/product="6K2"  
5714..6277  
/product="Nia-Vpg"  
6278..7009  
/product="Nia-Pro"  
7010..8566  
/product="N1b"  
8567..9367  
/product="CP"

BASE COUNT 2995 a 1818 c 2273 g 2612 t  
ORIGIN

Query Match 92.3% Score 788: DB 34: Length 9698:  
Best Local Similarity 96.1%: Pred. No. 0.00e+00:  
Matches 821: Conservative 0: Mismatches 33: Indels 0: Gaps 0:

Db 6117 AAGCTTTGAAGATTGACCTTAATGCCACACACCCACTCAAAAGTTTGTGCACAAAACAAATG 6176  
QY 1 AGCTTTGAGATTGATTGTATGCCATTAACCCACTCAAAATTTGTGCACAAAACAAATG 50  
Db 6177 GCAATGCAAAATTTCTCAGAGAGAGCTCGCACTAAGCAGAGCTGGCCAGCTGTAGAAG 6236  
QY 61 GCAATGCAAAATTTCTCAGAGAGAGCTCGCACTAAGCAGAGCTGGCCAGCTGTAGAAG 120  
Db 6237 TCAATGTGAAGACATACACGACAGAGGTGGAGCATGAAGCTAAATCGCTCATGAGAG 6296  
QY 121 TCAATGTGAAGACATACACGACAGAGGTGGAGCATGAAGCTAAATCGCTCATGAGAG 180  
Db 6297 GCTTGAGAGACTTCAATCCAAATGCCCCAAACAGTTTGTAGCTGAAAGTATCTGTTCAAT 6356  
QY 181 GCTTGAGAGACTTCAATCCAAATGCCCCAAACAGTTTGTAGCTGAAAGTATCTGTTCAAT 240  
Db 6357 ATGGGACATCAGAGTGTACGCTTTTGGATTGGAGCATACATATATACGGAACCAACCAAT 6416  
QY 241 ATGGGACATCAGAGTGTACGCTTTTGGATTGGAGCATACATATATACGGAACCAACCAAT 300  
Db 6417 TGTTTAGGAGTTACAAATGGTTCCATCGAGGTGCAATTCATCGAGGTACATTCAGGTGA 6476  
QY 301 TGTTTAGGAGTTACAAATGGTTCCATCGAGGTGCAATTCATCGAGGTACATTCAGGTGA 360  
Db 6477 AGAATCTACACACTTTGAGCTTCTGCCAATTAAGAGTAGGAGACATCATCTCTCATCAAAA 6536  
QY 361 AGAATCTACACACTTTGAGCTTCTGCCAATTAAGAGTAGGAGACATCATCTCTCATCAAAA 420  
Db 6537 TGCCGAAGAATTTCCCTGTCTTTCCACAGAAATTCGATTTCCGAGCTCCTATACAGATG 6596  
QY 421 TGCCGAAGAATTTCCCTGTCTTTCCACAGAAATTCGATTTCCGAGCTCCTATACAGATG 480  
Db 6597 AAAGAGTTTGTAGTTGGGACCAACTTTTCAGGAGAGTATGCATCGTCAATCAACACAG 6656











gene 1015..2409  
 /gene="HC"  
 mat\_peptide 1015..2409  
 /gene="HC"  
 gene 2410..3504  
 /product="helper component protein"  
 /gene="P3"  
 mat\_peptide 2410..3504  
 /gene="P3"  
 gene 3505..3660  
 /product="P3 protein"  
 /gene="6K1"  
 mat\_peptide 3505..3660  
 /gene="6K1"  
 gene 3661..5562  
 /product="6K1 protein"  
 /gene="CI"  
 mat\_peptide 3661..5562  
 /gene="CI"  
 gene 5563..5718  
 /product="6K2"  
 /gene="6K2"  
 mat\_peptide 5563..5718  
 /product="6K2 protein"  
 /gene="CI"  
 gene 5719..6282  
 /gene="VPg"  
 mat\_peptide 5719..6282  
 /gene="VPg"  
 gene 6283..7014  
 /product="genome linked protein"  
 /gene="Nia-Pro"  
 mat\_peptide 6283..7014  
 /gene="Nia-Pro"  
 gene 7015..8571  
 /product="Nia protease"  
 /gene="N1b"  
 mat\_peptide 7015..8571  
 /product="N1b protease"  
 /gene="CP"  
 mat\_peptide 8572..9372  
 /gene="CP"  
 3'UTR 9376..9701  
 /product="coat protein"  
 BASF CONT 3057 a 1755 c 2238 g 2651 t  
 ORIGIN

Query Match 62.1%; Score 530; DB 34; Length 9701;  
 Best Local Similarity 81.1%; Pred. No. 0.00e+00;  
 Matches 691; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

Db 5124 SCTCTAAAATGATTTGATGCCACACCAACCCACTCAAAATCTGTGATAAATCGAATGGC 6183  
 QY 3 GCTTTGAAGATGATTTGATGCCACATACCCACTCAAAATTTGTGACAAACAAATGGC 62  
 Db 5184 ATTGCTAAGTTTCCTGAAGAGAACTTCAGTTGAGGCAAACTGGCCAGCAATAGAGGTT 6243  
 QY 63 ATTGCCAATTTCTGAGAGAGAGTTCCAGTAAAGCCAGACTGGCCAGCTGTAGAGTC 122  
 Db 6244 GATGGAAGACATTTCCAAAACAGGAAGTGGAGCATGAAGCCAAATCACTCATCAGAGGT 6303  
 QY 123 SAGGTGAAGGACATACCAGCAGGAGGTGGAACATGAAGCTAAATCGCTCATCAGAGGC 182  
 Db 6304 TTAAGGGATTTCAATCCAAATTCCTCAACAGTTTCGACAGTAAAGTGTCTGTTGAATAT 6363  
 QY 183 TTGAGAGACTTCAAGCCAAATTCGCAACAGTTTGTAGGCTGAAGATATCTGTTGAATAT 242  
 Db 6364 GAAAGTCTGAAATATATGGTTCGGTTTGGTGTGATATATATATGAAGAAACCCATCTA 6423  
 QY 243 GGGACATCAGAGATGACGGTTTGGATTTGGAGCGTACATTAATAGGCAACCCACCTTTG 302

Db 6424 TTCAAGAGCTTCAATGGATCCATGGAAGTGGCATCAATGCGATGGAACATTTCAGAGTGAAG 6483  
 QY 303 TTCAGAGATTATAATGGTTTCCATGAGGAGGTACCATCGACGGTACATTCAGGTTAAAG 362  
 Db 6484 AATTTGCATAGCTTGAGGGTTTTACCGATCAAGGAGGAGACATTATCATATAAAGATG 6543  
 QY 363 AATCTACACAGTTTGAGCGTTCTGCCAATTAAGGTAGGGACATCATCTCATTAAGATG 422  
 Db 6544 CCAAGAGATTTCCCTGTTTTTCCCAAAAACATGCACTTCCGAGCTCCAGTCCAGAAATGAG 6603  
 QY 423 CCAAAAGATTTCCCTGTCTTTCCACAGAAATTCATTTCCGAGCTCCACAGAACGAA 482  
 Db 6604 AGGATTTGTTGGTGGAACTAATTTTCAAGAAAACATGCAATCATCATCATCACAGAA 6663  
 QY 483 AGAATTTGTTAGTTGGAACCAACTTTTCCAGGAGAGTATGCAATCATCATCACAGAA 542  
 Db 6664 ACGAGTACTACATCAATGTACCGGGCAGCAGCTTTTGGAGCATTTGGATTGAACAAAT 6723  
 QY 543 GCAAGCACTACTTACAATATACAGGAGCAGCATCTTCCGAGCATTTGGATTGAACAGAT 602  
 Db 6724 GATGGCATTTGGATTACCACTAGTAGTACAGCTGATGATGCTAGTGGATACAC 6783  
 QY 603 AATGGACACTGTGGACTACAGCTGTGAGCACTGCGAGTGGATGTCTAGTGGAAATTCAC 662  
 Db 6784 AGCTTGGCAATAATGTGCAAAACACGAAATTTATTATGAGCCTTTTATGAGGATTTTGA 6843  
 QY 663 AGTTTGGCAACAATGCACACACAGCACTACTACTACGCTTCGATGAGATTTTGA 722  
 Db 6844 AGTAGTATCTCCGAATTAATGAGCATATGAGTGGAGCAAAATCGTGGTATATTAACCA 6903  
 QY 723 AGCAAGTACTCCGAACCAATGAGCACAATGAATGGTCAAGTCTTGGATTATTAATCCA 782  
 Db 6904 GATACTGTGTTGGGTCCTTGAAGCTCAAGGAGAGTACCCCTTAAGGCTGTTTAAAG 6963  
 QY 783 GACACAGTGTGTTGGGCGCCGTGAAACTTAAAGACAGCACTCCCTAAAGGGTTATTTAA 842  
 Db 6964 ACAACAAAACCTT 6975  
 QY 843 ACAACAAAACCTT 854  
 RESULT 9  
 LOCUS PVYCI6K2 1784 bp DNA VRL 24-JAN-1994  
 DEFINITION Potato virus Y (strain O) CI, 6K2, VPg, N1a-Pro, N1a gene.  
 ACCESSION Z29526  
 NID g450360  
 KEYWORDS 6 kd protein; 6K2 gene; CI gene; Cytoplasmic inclusion protein;  
 genome-linked protein; n1a gene; polyprotein; protease; Sma.I  
 nuclear inclusion protein; VPg gene.  
 SOURCE Potato virus Y.  
 ORGANISM Potatoes.  
 Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;  
 Potyvirus.  
 REFERENCE 1 (bases 1 to 1784)  
 AUTHORS Wejnicky, M.A. and Baulcombe, D.C.  
 TITLE The nucleotide sequence of the central fragment of potato virus Y  
 (O strain) genome encoding 6K2 and VPg proteins  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1784)  
 AUTHORS Wejnicky, M.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-JAN-1994) Wejnicky M. A., Institute of Biochemistry  
 and Biophysics, Department of Protein Biosynthesis, ul. Pawinski 5A,  
 Warszawa, Poland, 02-106  
 FEATURES  
 source Location/Qualifiers  
 1..1784  
 /organism="Potato virus Y"  
 /strain="O"  
 /isolate="British"  
 /db\_xref="taxon:12216"  
 /clone="pPVY15"  
 1..850  
 gene



```
BASE COUNT      1016 a      560 c      788 g      894 t
ORIGIN
Query Match      38.4%      Score 328;      DB 34;      Length 3258;
Best Local Similarity 80.9%;      Pred. No. 5.68e-267;
Matches 429;      Conservative 0;      Mismatches 101;      Indels 0;      Gaps 0;

Db      1      ATGGAAGTGGCAATCAATGCATGGAACATTCAGAGTGAAGAATTTGCATAGCTTGAGTGT 60
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      324      ATGGAAGTACCATCCATGCAGGATACATTCAGGTAAGAATCTACACAGTTTGAGCGTT 393
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db      61      TTACGATCAAGGACAGACATTCATCATATAAAGATGCCAAGAGACTTCCCTGTTTTC 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      384      CTGCAATTAAGGTAGGACATCTCTCAATTAATAATGCCAAAAGATTTCCCTGTCTTT 443
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db      121      CCACAAAATGCACTTCGAGCTCCCAATCAAAATGAGAGATTGTTGTTGGTGAAC 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      444      CCACAGAAATGCAATTTCCGAGCTCCCTACACAGAACGAAGAATTTGTTAGTTGGAACC 503
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db      181      AACTTTCAGAAACAAATCGCTGCTCAATCATCACAGAAACGAGTACTACATCAATGTG 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      504      AACTTTCAGGAGAGATGATGATGCTGATCATCACAGAACGACACTACTACATATA 563
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db      241      CCGGCACACACTTTTGGACGATGATGATGAAACAAATGATGGGCAATGTGGATACCA 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      564      CAGGACGACATCTTGAAGCATTTGGATTTGAAACAGATATGACACACTGTGGACTACCA 623
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db      301      GTAGTGATGACGATGATGATGCTGTTGGAATACACAGCTTGGCGAATATGTGCAA 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      624      GTGTGACATGCGGATGGAATGCTAGTGGGAATTCAGATTTGGCAACAAATGACAC 683
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db      361      ACCAGCAATTAATTAATGCTTTGATGAGGATTTGAAAGCAAGTACTCTCCGAACATA 420
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      684      ACCAGCAATTAATTAATGCTTTGATGAGGATTTGAAAGCAAGTACTCTCCGAACATA 743
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db      421      GAACACATGATGAGCAACAAATCGTGGTATATACCCAGATATCTGTGTGGGGTCCA 480
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      744      GAGCAACATGATGAGTCAAGCTTGTGATTTATATACGACACAGTGTGTGGGGCCG 803
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db      481      TTGAAGTCAAGGAGTACCCCAAGGCTGTGTTTGAAGCAACAAACT 530
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      804      TTGAACATTAAGACAGCACTCTTAAGGGTTATTTAAACAAACAAAGCT 853
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT      11
LOCUS      PMVCG      9640 bp ss-RNA
DEFINITION      Pepper mottle virus complete genome.
ACCESSION      M96425
NID      932869
KEYWORDS      coat protein; complete genome; helicase; inclusion body protein Ci;
      inclusion body protein N1a; inclusion body protein N1b; nuclear
      inclusion protein; pro-1 protease; pro-3 protease; replicase.
SOURCE      Pepper mottle virus (strain California) cDNA to genomic RNA.
ORGANISM      Pepper mottle virus
      Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
      potyvirus.
REFERENCE      1 (bases 1 to 9640)
AUTHORS      Vance,V.B., Moore,D., Turpen,T.H., Bracker,A. and Hollowell,V.C.
TITLE      The complete nucleotide sequence of pepper mottle virus genomic
      RNA: Comparison of the encoded polyprotein with those of other
      sequenced potyviruses
JOURNAL      Virology 191, 19-30 (1992)
MEDLINE      93033110
FEATURES      Location/Qualifiers
      source      1..9640
      /organism="Pepper mottle virus"
      /strain="California"
      /db_xref="taxon:12209"
      1..167
      168..1028
      /product="pro-1 protease"
      168..9374
      5'UTR
      mat_peptide
      CDS
```

```
/note="orf1"
/codon_start=1
/db_xref="PID:932870"
/translation="MAISVIOFGSEFVNLPKSPQCTTTHCHPCPKSMSTINVRPDDPFA
ELEKLEPEYLOKMDATIRQTKGTLYNKHMSAKRKLRKKKQREHEEVRLEPNAFA
YIVSNTITGGSEVSKMEVEIKRPLTKSRKIKSLTPTVFDGHMRELFELRDC
ATRSMVHLIGKRTKLEAFKRRASNAVATLHMRGVRKRDIVLEENWNVLSNL
SKVTWGLFHAESLTKGDSGLILNARALPGKRCGRGFEIVRGKSGVVLDAESKL
SMATVHMEQSTPEACWSGLEKWSVRKPTAHTCKPTYSVSCNGEVAALIAAALF
CHKLTGCECSKEICDLTSNECVQELYKNTSLALERMNHLHPEFHILVKVLSVRLIE
ASNGHTEIFDEIFKMGSTQSPFTHLNKNEFMKGNENSTGEWLTAQHLRELFH
QKNTDNIKKGLDASFRNSARQAYNLISCDNOLDKNASFLMGQREYHARRFLNF
FQIDPSKGLAYEDFTTPNGSRKALNLTIVPLDLAFAFRPMNG:ITQPPYKGYCT
SOLDGNEVYPCCTTLDGQPIRSAYVAPTKKHLVNGTGDYKINLPGKGTETLYTA
LDGVCYINILAMLVNISEEAKDTKKVRDIEPKLWKPFLMDLATTCAQLRIFHP
DVHDAELPRILLVDHNTQCHVVDYSGISGTHYLKAATVSQVLSVLFADDDNLESIKHY
RVGIVENHKVQIDNQPSRCVSEFHAIKMLIKGIRYPSVMYELLSEEPYLLVLSLS
PSIIIAMYNDRAFELAVOILEKEOSIPLATILTNLAAYSVATITVOOQLLELSA
DOLLNVTGDFRVFAYQSALTLLTRMDQAKANSELSGGFNEYDQDLAWTLEKQF
GLLDHDKWLSLEKFRYWSRRKRLRSNIKSRSPVASALSLSLKFPMGKVS
HMKAGVCTKQGTNFDARCLGISTVFGSLMRKFPKSAKVLSSLVGLGALLNTHA
ANRLLDNPISEHAALLEYKEDTCHELYTALERKLGKAPTWDVSCYVAK.NPAM
LEF.KDSYDEKQVHQRSTIDLKVEHLIAFVLAIMLFDSESDCVKTLNKGKVV
CSLGSVGRHQSLDDFVSTMDKNEFVDFELNDSVRKNLATTETTFESWDEQVARGT
IPHYTEGRFMETRATAKAVASDISISSEPDFLIPGAVSGKSTGLPHLSTYGRVL
LIEPTPLAENFVKQLSGPFFLPTMRMGRNSVFGSPISVMTSGFALHPFANNITQ
LQETOFIIDECHVMDASSMAFRSLIHTYHTNCKLVKVSATPPGQREVEFTTQFVKLV
VEDLSKFTVESOGTSNCDMIOYGNLLVYVASYNEVDOLSKLLVARENVTKVDS
RTMKHGEIIVRTGKSPHFVVAINTIENGVTLDIVDIFGKMSFPFLDNDKRSVA
YKHYSEIGERIORLGRVROKGTALRIHTEKGLIEIPOMISTEALCYFAYNLIV
MSSGTSPIRKNCCTIPQVTHTEFELSPFMNEVSHDGTMPVYVHETLKRYKLRDVS
PLGESSIPYRASSDITAGDYRPGVKLIDIPDETIFAHLKTFHRKFTNNLWESVLK
YKGAAPPTLRSSSIIKIATYTLSDLVAPRTLAUVESLLEDERTKQYQKSLDNGC
SSMSEVGIISNALRAKYSKDHTEINKLETVKQALKEHNLGSGDELNLKFEESL
QFVHQSSSLAKALGLGVWKNLSLIVDAILAAGVACGGAWLLTFTAKMSEVSHC
GRSTKRIQALFKRKARDKRAFGEIDNDEETIEYFSGATYKKGKGTITVGMGRINR
FINNYGEPPQFSYIKFVDPITGAQNEENVYADIVQENFGDIPROMILDDLDRP
QTDVHTIYAILIKDWSNALKVDLTPHNPLRVSQKASAIKMKPERGELQTOQAVE
VDVDTIPKEVYVKEATLMPGLPDYPIAQTQVKI.TYKSELGETSYGLGFGGLIIAN
HHLKFSNGSLVSKHGVFRVPLMAISVLPKGROMI.IIKPKDPPFPVQKFRF
PASTDVLGLISNQRYISTITVSEISATVPVPRSTFWKHWTIDDSHGCLPIVSTID
FGILGLSLANNRSENSTYITAFDSFMKILSRSGENTEWKNWKNYINPDIVLWGLQLT
KGTGPMFKITTKMIEDLAFKESVREGAHTSSWMLKYLKLNKLAIAIMKSLQVTKHY
VKGECMFQYLOENPRANEPQPMWAGKSMLNKEAYIKDIMKYSKIDVGVVDCD
PHIRKLSLELLYTOIHGPKGYSITDEEIEFKALNITTAGVAMYGKKKVEFEKFTIE
DKAETLRQSLPLTKLGVMEWALKAEKRSKEIKANKTTRTFTAAPDITLLGKQVY
DDLNNQYSKNTECQWMTKYGGWPKLLTALPAGWIYCDAGSQDSSSLTPLYLN
AVLTIRYAFMEDWDIGYKMLQNLTYETIITPDGTIIVKFRGNSSQGSTVVDNS
LMVLAMHYAFVREGIAPEEIDSCIKFEVNGDULLIAVNPRESLDDTLNSHFSLSL
NYDESRTPNKSSELWMFSGHSGISVEGTYIPKLEERIVSILOWDRAETPEYRLAICA
AMIESWGPQLTHIEIRFREYSWLEKNPYADLASEGKAPYISELAKLLYNOQVQMS
FRSYLYFADADEBEFCGTVEYRHQSSRSOTLDAGEBKKNKEVATVSDGMGKVEE
STROSDVNGATVGTFTPIKISITEKMBPKOKSGVLNLAHLHLPKSDQNDIENRS
TQAQFDNMYCEWKAAYDQEAEMGTVMNGLMWCIENGTSINISGTTWMDQCEQVEF
PKUPVIENAKPTFOIMAHFSDVAEAYIEMRNKQPEYMPYGLVNRNMDGLARYAFD
FYEVTSRTSTFAPEAHIQMKAAALKSAQTPLFGDGGIGTGIGTERTHTEDVSPDMH
TLLGVREM"
1029..2396
/product="helper component-protease"
2397..3635
/product="pro-3 protease"
3636..5537
/product="inclusion body protein Ci"
5538..5693
/product="6kD protein"
5694..5995
/product="inclusion body protein N1a"
5996..8552
/product="inclusion body protein N1b"
8553..9371
/product="coat protein"
9375..9640
3'UTR
BASE COUNT      3051 a      1747 c      2235 g      2607 t
```

[illegible]









QY	567	GGCAGCACATTCTGGAAGCATTGGATTGAAACAGATAATGGACACTGTGGACTACCAJTG	625
Db	6759	GTGAGCAGCTCGAGATGGCAGTATCTTGGGTGTACACAGCTTTGCAAAATTCACAGCAACACC	6818
QY	627	GTGAGCAGCTGCGGATGGATGTAGTCGGAATTCACAGTTTGGCAACAATGCACACACC	686
Db	6819	CAGAATTTCTATCGAGCTTTCCCTGACAACTTCGAGAGCACCACATCTTGTCAAAATCAAGAC	6878
QY	687	AGCACTACTACTAGCCTTCGATGAAGATTITGAACGAAGTACTCTCGAACCAATGAG	746
Db	6879	AATGATACTGGATAAAGCAGTGGCGATACAAATCGGATGAAGTTTGTGGGG	6931
QY	747	CACAATGAATGGTCAAGCTTTGGATTATATCAGACACAGCTTGTGGGG	799
RESULT 15			
LOCUS	PPVNIAPRO	1695 bp	mRNA VRI 03-AFF-1992
DEFINITION	Plum pox potyvirus ORF containing NI-a-like protein mRNA, partial cds.		
ACCESSION	M26965		
NID	g333303		
KEYWORDS	NI-a-like protein; cleavage protein.		
SOURCE	Plum pox potyvirus (strain Rankovic) cDNA to mRNA.		
ORGANISM	Plum pox virus		
REFERENCE	Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae: Potyvirus.		
AUTHORS	1 (bases 1 to 1695)		
TITLE	Garcia,J.A., Riechmann,J.L. and Lain,S.		
JOURNAL	Proteolytic activity of the plum pox potyvirus NI-a-like protein in Escherichia coli		
MEDLINE	Virology 170, 362-369 (1989)		
FEATURES	Location/Qualifiers		
source	1..1695		
CDS	/organism="Plum pox virus" /strain="Rankovic" /db_xref="taxon:12211" 1..1695 /partial /note="ORF" /codon_start=1 /db_xref="PID:g333303" /translation="FKNLIDPSYELVNFPALECVHQTKEGVSKALQIKGHNKP LITRDATLMLGVGGWAMIFSYLRDSKEEVDHESKSLPGLRDYNPDIASSICOLNNS VYGDSTMDVYEGSAYSKGKTRGMGTTRKFNMYGVDPTDYNVFPDPLTG HTLDENPLMDINLQVEHFSQIRNDYIGDDKITMHOIMNSPGIVAVYIKDATQKALVD LTPHNPRLVCDKTATAGPPEFERPLMGAYOKSALNKDAYVKD FRGURYNPIASSICOLNNSGAGSVMFGGLGLVITNOHLKFRNDGELTIRSHHG EFVVKDTKLKLPCKGRDIVITRPKDPPFPKRFQRTPTEDRVCLIGSNFQTS ISSTKSETSATVPDNSHFVKHWISLKDGHGLPIVSTRDGSILGLHSLSANTQNF YAFDPDNFTTLYSQNDNWMIKWRYPDEVCKWSLQKRDIPQSPFTICKLLTDLG DFVYQSTQTHWLDRLEGKLVKAVGACPGQLVTKHVVGKCTLFETVLLTHPEHE" 238..1545 /gene="NI-a-like protein" 238..1545 /gene="NI-a-like protein" /product="NI-a-like protein" 539 a 355 c 391 g 410 t		
BASE COUNT	539 a	355 c	391 g 410 t
ORIGIN			
Query Match 13.8%; Score 118; DB 34; Length 1695;			
Best Local Similarity 58.8%; Pred. No. 2.73e-75;			
Matches 393; Conservative 0; Mismatches 275; Indels 0; Gaps 0;			
Db	802	GAAGTTGACACAGTAAGTAAATCAGTTCAGAGCCTGAGAGACTATAATCCATCGCA	861
QY	147	GAGGTGGAACATGAAGCTAAATCGCTCATGAGAGCCTTGAAGAGACTTCAACCTATTGC	206
Db	862	AGCTCAATATGCAATTTGAATAACTCATCTGCTGTAGACAAAGTGAATGTTGGACTT	921
QY	207	CAACAGTTTGTAGCTGGAAGTATCTGTGTAATATGGACATCAGAGATGACGTTT	266

LEXVGDENPDLLKHAEDLIGDGVVHVHVKSRDQANLERNVAFVALVMMFLDSESDG VYKLLKNGIMSGSVRAVHQSLDDIEDLDEKLVTVDFVLSQSNVAVTPFDDSTFE KMMNOLKCNVLPVPHVTECHLEFTRANAHTANEMVHGSODLILCAVGSKSTG LPNLSKXKGVLLIETPRPLAENVCKQLRQCPENVPILMRGMSTFGSTPIVWTS YALHFLANNVTLDNKCILFDCSHVDASAMAFRCLLSEYSTPLKLVKATPGPHE VERTKQEVKVIIEESLSFOFNSDILKHSNVLVYVASYNEVDLTKLL TDRSKVYDGTMRKMGVNEIPTSGTQAKPHEVATNIENGVTLDIDVDFGLKV VPVLDIDNLRVYTKISISYGERIQRLGRVKNPGAAIRIGTFEKLQIPIIATE AAFLCTYGVSEVNTSVLLAMCVKQARTMOQFELSPFVTVAVLRFDGMHOFIF RLLSKVRDSEVILANKLAIPNSVGVMSVDRYKROGNLDDLDERVFPVKKDPE TLHERIWOAVETHKSDAGFRIGSSACKIAYILQTDIHSIPRTIKIIDLLEQERTK QAFRMTSQCSSPSSLSITSAIRSKYAKDHTTEENIGVLOMAKSQULEPKNUNID PSYPELIRNFALGECVHQHTEKGVSKALQKGNKRLITRQATLMGLVGGWAMIF SYLRDSKEEVIHQGNRRQOKLFRQARDNRMAREVYCDSTMEAYFGSAYSKGK SKGTRMGKTRKFRVNMGYDPTDFVRFVDPDLTCHTLDSEPLMDINLVOEHSOI RNDYIGDKITMHOIMNSPGIVAVYIKDATQKALVDLPHNPLRVCDKTATAGPPE REFELROTGHVFEVPEPNAIPKINEEGDEVDHESKSLPGLRDYNPDIASSICOLNNS CARQSEMFGLGGLGLVITNOHLKFRNDGELTIRSHGEFVVDKTLKULPCQGRDIV IIRLPDPPFPFRRLOFRTPTEDRVCLIGSNFQTSISSTKSETSATVPDNSHFVK HWISLKDGHGLPIVSTRDGSILGLHSLSANTQNFYAFDPDNFTTLYSQNDNWM IKQWRYPDEVCKWSLQKRDIPQSPFTICKLLTDLGDFVYQSTQTHWLDRLEGK LKAVGACPGOLVTKHVVGKCTLFETVLLTHPEHEFERPLMGAYOKSALNKDAYVKD LMKSKPIVVGAVDQOFERAVDVVLSMLISKGFECNVVTPDDDIFSLNKKAAVGA LYGSKRDYFENVSDOKSEFVASCRLFRFGKGVWNGSLKRAELRPKEKVEANKTRS FTAAPIDTLGKVCYVDFNNQFYSNLHCPWVGMTKFRGWDKLLKALPEGVYCD ADGSQDSSLSPLYLINAVLIRLAFMEEMDIGQMLSNLYTEIVYPIATPGLTVYK FKGNNSQPSVVDNTLMVILAMTYSLKLGHPDTHDGCIRYFVNGDLDLAVHPAY ESYDEQHEFSOLGNTYTKTENKEELFWMSHGVLVDOMYIPKLEPERIVSTLE WORSNEPILRELAECASMEAWGYKELLREIKFYSKVLQEAQYNALSKDGAPYIAE TALKLYIDTEASETEIERLEAFYDINDDGESSNVVHQADERDEEDVADALQPPVY IQAPRTTAPMLPIFTPIATQATPKVSQVSPQLOTFGTFSHEDASPSNVALNT NORVDVAGSTGTFVPLKAMTSKSLPKVKGKAIMNLHLAHYSPADVLSNTRAP QSCFTWYEGKRDYDVTDEMIIILGLMWCIEINGTSPNNGMMVMDGTQVEYEP IKPLDHAKPTFOIMAHFSNVAEVIKRNVEKAYMPRYGIORNLTDYSLARYAFDF VEMTSTTPVAREAHIQMKAALRNQVNRILFGLDGNVGTQBEDTERHTAGDYNRNHN LLCMRGV"	QY 567	GGCAGCACATTCTGGAAGCATTGGATTGAACACAGATAATGGACACTGTGGACTACCAJTG	625
Db	6759	GTGAGCAGCTCGAGATGGCAGTATCTTGGGTGTACACAGCTTTGCAAAATTCACAGCAACACC	6818
QY	627	GTGAGCAGCTGCGGATGGATGTAGTCGGAATTCACAGTTTGGCAACAATGCACACACC	686
Db	6819	CAGAATTTCTATCGAGCTTTCCCTGACAACTTCGAGAGCACCACATCTTGTCAAAATCAAGAC	6878
QY	687	AGCACTACTACTAGCCTTCGATGAAGATTITGAACGAAGTACTCTCGAACCAATGAG	746
Db	6879	AATGATACTGGATAAAGCAGTGGCGATACAAATCGGATGAAGTTTGTGGGG	6931
QY	747	CACAATGAATGGTCAAGCTTTGGATTATATCAGACACAGCTTGTGGGG	799
RESULT 15			
LOCUS	PPVNIAPRO	1695 bp	mRNA VRI 03-AFF-1992
DEFINITION	Plum pox potyvirus ORF containing NI-a-like protein mRNA, partial cds.		
ACCESSION	M26965		
NID	g333303		
KEYWORDS	NI-a-like protein; cleavage protein.		
SOURCE	Plum pox potyvirus (strain Rankovic) cDNA to mRNA.		
ORGANISM	Plum pox virus		
REFERENCE	Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae: Potyvirus.		
AUTHORS	1 (bases 1 to 1695)		
TITLE	Garcia,J.A., Riechmann,J.L. and Lain,S.		
JOURNAL	Proteolytic activity of the plum pox potyvirus NI-a-like protein in Escherichia coli		
MEDLINE	Virology 170, 362-369 (1989)		
FEATURES	Location/Qualifiers		
source	1..1695		
CDS	/organism="Plum pox virus" /strain="Rankovic" /db_xref="taxon:12211" 1..1695 /partial /note="ORF" /codon_start=1 /db_xref="PID:g333303" /translation="FKNLIDPSYELVNFPALECVHQTKEGVSKALQIKGHNKP LITRDATLMLGVGGWAMIFSYLRDSKEEVDHESKSLPGLRDYNPDIASSICOLNNS VYGDSTMDVYEGSAYSKGKTRGMGTTRKFNMYGVDPTDYNVFPDPLTG HTLDENPLMDINLQVEHFSQIRNDYIGDDKITMHOIMNSPGIVAVYIKDATQKALVD LTPHNPRLVCDKTATAGPPEFERPLMGAYOKSALNKDAYVKD FRGURYNPIASSICOLNNSGAGSVMFGGLGLVITNOHLKFRNDGELTIRSHHG EFVVKDTKLKLPCKGRDIVITRPKDPPFPKRFQRTPTEDRVCLIGSNFQTS ISSTKSETSATVPDNSHFVKHWISLKDGHGLPIVSTRDGSILGLHSLSANTQNF YAFDPDNFTTLYSQNDNWMIKWRYPDEVCKWSLQKRDIPQSPFTICKLLTDLG DFVYQSTQTHWLDRLEGKLVKAVGACPGQLVTKHVVGKCTLFETVLLTHPEHE" 238..1545 /gene="NI-a-like protein" 238..1545 /gene="NI-a-like protein" /product="NI-a-like protein" 539 a 355 c 391 g 410 t		
BASE COUNT	539 a	355 c	391 g 410 t
ORIGIN			
Query Match 13.9%; Score 119; DB 34; Length 9741;			
Best Local Similarity 59.1%; Pred. No. 3.62e-76;			
Matches 386; Conservative 0; Mismatches 267; Indels 0; Gaps 0;			
Db	6279	GAAGTTGACACAGTAAGTAAATCAGTTCAGAGCCTGAGAGACTACAATCCATCGCA	6338
QY	147	GAGGTGGAACATGAAGCTAAATCGCTCATGAGAGCCTTGAGAGACTTCAACCCAATTGCC	206
Db	5339	AGCTCGATATGCCAATTGAATAACTCATCTGGTGTAGACAAAAGTGAATTTTGGACTT	6398
QY	207	CAACAGTTTGTAGCTGAAAGTATCTGTGTAATGGACATCAGAGATGTACGGTTTT	266
Db	6399	GGCTTTGGGGTTTAATGTCAGGAATCAGATTGTTTCAAAAGGAATGACGGAGAGCTA	6458
QY	267	GGATTTGGAGCGTACATAATACGGAACCACTTGTTCAGGAGTTTAATGTTTCCATG	326
Db	6459	ACAATCCGATCGCATCATCGGGAAATTCGTAGTGAAGGACACAAAACCTCTCAAACTGCTT	6518
QY	327	GAGGTACGATCCATCGACGGTATCTCAGGGTAAAGATCTTACACAGTTTGAGCGTTCTG	386
Db	6519	CCTTGCAGAGTGGGACATAGTATCATCAGATTACCAAGGACTTCCCTCTCTTTCCG	6578
QY	387	CCAAATTAAGGTAGGAGCATCATCTCATTAATAATGCCAAAAGATTTTCCTGTCTTCCA	446
Db	6579	AGGAGTTTGCAGTTCCGACCCCGACGACTGAGGACAGCTTTGTTTAATTGGTTCAAAAT	6638
QY	447	CAGAAATTCGATTTCCGAGCTCCTACACAGACAGCAAGAAATTTGTTAGTTGGAACCAAC	506
Db	6639	TTCCAAACGAAGAGCATTTTCAAGCACCATGTGCGGAACCAAGCACAACATATCCAGTTGAT	6698
QY	507	TTTCAGGAGAAGTATGTCATCGTCGATCATCACAGAACCAAGCACTACTTACAAATATACCA	566
Db	6699	ACAGTCTATTTCTGGAACACTGGATTAGTACGAAGGATGGTCATTGCGGATACCCCATC	6758





























[illegible]











































pPR1, pPI coexpression	0.00e+00
pAct-Splice	n 0.0e-00
Plasmid pF73 encodin	0.00e+00
PAPC-CW1 expression v	0.00e+00
PAN-CW1 contg a Tum	n 0.0e-00
Plasmid PMV1 contain	0.00e+00
Plasmid PKSEL5	0.00e+00
Sequence of clone pS2	0.00e+00
Plasmid GS contg. N.	0.00e+00
Plasmid glucosylase	0.00e+00
Plasmid pPBPAT (rat	0.00e+00
PAN-CW1	0.00e+00
Plasmid pBSGFP expres	0.00e+00
Sequence of clone pS2	0.00e+00
Plasmid pMVI	0.00e+00
Sequence of clone pF2	0.00e+00
Plasmid pM6-1	0.00e+00
Plasmid pM6	0.00e+00
Plasmid pSVFL(-) inc	0.00e+00
vector pF7M1-lor	0.00e+00
Plasmid pWP3196-enco	0.00e+00
Construct pSEM-HTR(K)	0.00e+00
Construct pSEM-HTR CO	0.00e+00
Human cyclin D1-human	0.00e+00
Shortened C1 gene in	0.00e+00
pBD10	0.00e+00
pDBL0; plasmid LNA i	0.00e+00
Plasmid pVEI08 used i	0.00e+00
pVEI08	0.00e+00
pD54; plasmid RNA rep	0.00e+00
Sequence of pTE4	0.00e+00

## ALIGNMENTS

RESULT 1  
ID T43137 standard: DNA: 5534 BP.

10-FEB-1997 (first entry)  
pUMIGIT sequence including upstream activating sequence.  
transgenic plant, gene expression: upstream activating sequence:  
UAS: transactivating protein: Gal4; herbicide resistance:  
polyhydroxybutyrate: safety: pUMIGIT; ds.

OS Synthetic.  
PN CR1350030-A.  
PD 09-AUG-1996.  
PP 150039.  
PR 24-MAY-1995; GB-002456.  
PR 08-FEB-1995; GB-002456.  
PA (UYWA) UNIV WARWICK  
PA Bennett M, May S, Ramsay N:  
WPI: 96-486150/49.  
DR Control of genes in transgenic plants - using an upstream activating  
PR sequence activated by a transactivating protein expressed using a  
PT separate promoter.  
PT Example: Fig 5F: 48pp; English.  
PT Reporter plasmid pMUGIT (T43137), or pUAS Minimal promoter Gus  
CC Int Terminator, contains a beta-glucuronidase (GUS) reporter  
CC gene under the control of the 46S cauliflower-mosaic virus minimal  
CC promoter and 10 synthetic 17-bp GAL4 binding sites. The upstream  
CC activating sequence (UAS) of pMUGIT is activatable by yeast  
CC transactivating protein GAL4. In a novel method for controlling  
CC gene expression, a first transgenic plant carrying a gene encoding  
CC a desired phenotype (herbicide resistance of polyhydroxybutyrate  
CC prodn.) operatively linked to a UAS recognition site is pollinated  
CC by a second transgenic plant carrying GAL4 DNA (see also 1431350).  
CC The transgene is fully expressed in F1 hybrid plants but segregates  
CC apart in subsequent generations, improving environmental safety.  
SU Sequence 5534 Bp: 1420 A: 1355 C: 1387 G: 1363 T:

Query Match	84.9%	Score 2855	DE 25	Length 5534
Best Local Similarity	100.0%	Pred. No. 0.00e+00		

Q

```
*****
              (TM)
*****
Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

*****
MPsrchLun   n.d. - n.a. database search, using Smith-Waterman algorithm

Run on:      Sun Mar 14 09:18:30 1999:    WasPar time 487.74 seconds
           938.093 Million cell updates/sec

Tabular output not generated.

Title:       >US-09-020-716-1
Description:  (1-3363) from US09020716 seq
Perfect Score: 3363 1 TCACATCTCAAGGCGGGTGTTC... GCGGTGTGTGTGTGCCCATGG 3363
N A Sequence: N A Sequence: ... CCGGAACGACCACCCGGTACC
Comp.

Scoring table: TABLE default
Gap v
```

Post-processing: Minimum Match 08

[illegible]

statistics:  
Mean 10.015; Variance 5.930; scale 1.689

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		ID	Description	Pred. No.
		Match	Length			
C 1	2855	84	9	5534_25	pUMIGIT sequence incl	0.00e+00
C 2	2824	84	9	V32042	Plasmid pWP3169 encO	0.00e+00
C 3	2820	83	9	5585_14	Plasmid pMP6-IL2	0.00e+00
C 4	2818	83	8	5585_29	Interleukin-2 express	0.00e+00
C 5	2774	82	5	6305_6	K lactiSUS cerevisiae	0.00e+00
C 6	2629	75	2	2937_24	Plasmid pTRP.	0.00e+00
C 7	2529	75	2	2937_24	Plasmid pTRP.	0.00e+00
C 8	2395	71	2	5248_24	Vector pET Trc S05/N1	0.00e+00
C 9	2395	71	2	5248_17	plasmid pET Trc S05/N	0.00e+00
C 10	2390	70	8	3664_11	plasmid pASK46 for ex	0.00e+00
C 11	2375	70	6	6714_13	pWB254b plasmid expre	0.00e+00
C 12	2256	67	1	3832_11	Plasmid pASK676-Strp	0.00e+00
C 13	2254	67	0	4515_8	pLISC-SE vector	0.00e+00



















PI which allows stable multiple integration of DNA for prodn. of

PI heterologous proteins  
 PS Claim 1, Fig 1, 26pp; English.  
 CC This sequence represents a genetic vector which allows the stable  
 CC multiple integration of DNA sequences into the genome of Kluyveromyces  
 CC lactis and Saccharomyces cerevisiae. This sequence can be used in an  
 CC integrating vector which comprises a region necessary for the stable  
 CC maintenance of the plasmid in E. coli and a domain which acts as an  
 CC integrating unit consisting of two not contiguous sequences of the 255  
 CC ribosomal DNA from S. cerevisiae, flanking a genetic marker suitable  
 CC for selection of the yeast transformants in which the integration  
 CC event has occurred. Other DNA sequences may be introduced into the  
 CC integration plasmid, such as expression cassettes. The gene Hts3  
 CC from K. lactis and S. cerevisiae is pref. used as a genetic marker  
 CC for the selection of transformants and an expression cassette for  
 CC production and secretion into the culture medium of human lysozyme.  
 CC This complete transformation vector is 7850 bp long and includes the  
 CC integration vector of the invention and an expression cassette  
 CC comprising the K. lactis GAL7 promoter, the signal sequence of the K.  
 CC lactis killer toxin, the cDNA encoding the ripe form of human lysozyme  
 CC (Ht2) and the transcription termination signal FLP of the 2 micron  
 CC plasmid from S. cerevisiae.  
 CC Sequence 6824 BP: 1915 A: 1521 G: 1726 C: 1762 T:

Query Match 92.5% Score 2774; DB 6; Length 6824;

Best local similarity 99.3% Pred Nc 0.00e+00;

Matches 2940; Conservative 0; Mismatches 12; Indels 9; Gaps 3.

Db 3896 ccaattgcgcctataatgagctgattacaat---tcaetggccgtcggtttacaagtc 3952  
 Cp 2461 ccaattgcgcctataatgagctgattacaat---tcaetggccgtcggtttacaagtc 2821  
 Db 3953 gtgaattgggaacacccctggcgtttaccccaacttaactgcgcttgcagcacatccctttcg 4012  
 Cp 2820 gtgaattgggaacacccctggcgtttaccccaacttaactgcgcttgcagcacatccctttcg 2761  
 Db 4013 cgaagctgaatgaatgagcgaag 4072  
 Cp 2760 cgaagctgaatgaatgagcgaag 2701  
 Db 4073 tgaattggcgaatgagcgaag 4132  
 Cp 2700 tgaattggcgaatgagcgaag 2644  
 Db 4133 ttacgcgaacgtgacgcgttacacttgcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 4192  
 Cp 2643 ttacgcgaacgtgacgcgttacacttgcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2584  
 Db 4193 tcccttccctttcttgcagcgttgcagcgttgcagcgttgcagcgttgcagcgttgcagcgttgc 4252  
 Cp 2583 tcccttccctttcttgcagcgttgcagcgttgcagcgttgcagcgttgcagcgttgcagcgttgc 2524  
 Db 4253 cttaggggttccgattagtggtttacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 4312  
 Cp 2523 cttaggggttccgattagtggtttacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2464  
 Db 4312 atggttcaactagtgagc 4372  
 Cp 2463 atggttcaactagtgagc 2404  
 Db 4373 cccgcttctttaaagtgagctctgttccaaactgagagagagagagagagagagagagagagag 4432  
 Cp 2403 cccgcttctttaaagtgagctctgttccaaactgagagagagagagagagagagagagagag 2344  
 Db 4433 tctattctttgattataagggattttgcgatttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 4492  
 Cp 2343 tctattctttgattataagggattttgcgatttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2284  
 Db 4493 tgaatttaacaaatttaacgcgaatttttaacaaabattataacggtttacaatttccagg 4552  
 Cp 2283 tgaatttaacaaatttaacgcgaatttttaacaaabattataacggtttacaatttccagg 2227  
 Db 4553 tgaattctttggggaatgagc 4612

Cp 2229 tggcagcttttgcggggaatgtgtggggaacacccctatttttttttttttttttttttttttttt 2167  
 Db 4613 aaatatatccgcctcatgagacaataaacccctataataatgcttcaataatatttaaaaaa 4672  
 Cp 2166 aaataatgctgcttcaigagacacaaataaaccttgaataatgttcaataatatttaaaaaa 2107  
 Db 4673 gaagagatagatgatttaaatcttccatgtgcgccttattcccttttttttttttttttttttt 4732  
 Cp 2106 gaagagatagatgatttaaatcttccatgtgcgccttattcccttttttttttttttttttt 2147  
 Db 4733 ccttcctgtttttgctccaccagagacgcgtggggaagagagagagagagagagagagagagag 4792  
 Cp 2046 ccttcctgtttttgctccaccagagacgcgtggggaagagagagagagagagagagagagag 1987  
 Db 4793 ggggc 4852  
 Cp 1986 ggggc 1927  
 Db 4853 tgc 4912  
 Cp 1926 tgc 1867  
 Db 4913 attatccgttattgacgc 4972  
 Cp 1866 attatccgttattgacgc 1807  
 Db 4973 tgaattggctgaatgacgc 5032  
 Cp 1806 tgaattggctgaatgacgc 1747  
 Db 5033 agaattatgagctgctgccataaccatgagtgataacactgcgcgcgcgcgcgcgcgcgcgcgc 5092  
 Cp 1746 agaattatgagctgctgccataaccatgagtgataacactgcgcgcgcgcgcgcgcgcgcgc 1687  
 Db 5093 aacgatcgc 5152  
 Cp 1686 aacgatcgc 1627  
 Db 5153 tgc 5212  
 Cp 1626 tgc 1567  
 Db 5212 cagctgc 5272  
 Cp 1566 cagctgc 1507  
 Db 5273 tctagcttccgc 5332  
 Cp 1506 tctagcttccgc 1447  
 Db 5333 tctagcttccgc 5392  
 Cp 1446 tctagcttccgc 1387  
 Db 5393 tgggtctgc 5452  
 Cp 1386 tgggtctgc 1327  
 Db 5453 tatctacgc 5512  
 Cp 1326 tatctacgc 1267  
 Db 5513 aggtgcctcactgattgaagcgttgc 5572  
 Cp 1266 aggtgcctcactgattgaagcgttgc 1207  
 Db 5573 gattgatttaaaacttctattttaaattaaagagatctagtgagatccttttttttttttttt 5632  
 Cp 1206 gattgatttaaaacttctattttaaattaaagagatctagtgagatccttttttttttttttt 1147  
 Db 5633 tctcatgacccaaatcccttaacgtgattttcttccactgacgcgcgcgcgcgcgcgcgcgcgcgc 5692





































(I)

```

MPsrch_n0 n a - n a database search, using Smith-Waterman algorithm
Run on: Sun Mar 14 07:41:40 1999. Maspar time 4281.65 Seconds
1406.556 Million Cell updates/sec

```

```
>US-09-020-716-1
Description: (1-3363) from US00020716 seq
Perfect Score: 3363
N.A. Sequence: 1 TCGACCTCCGAGGGGATGAGCCTC
Comp.: AGTGTAACTTCCTGTTGCTGCCTGCCCATGG 3363
          GGGTAAGGACCACTTGATTACC
```

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0: Query 0

Searches:

Post-processing: Minimum Match 08  
Listing first 45 summaries

Database: emb1-est56  
1:em est1 2:em est1 3:em est2 4:em est3

Database:

5:gb_est1	6:gb_est1c	7:gb_est11	8:gb_est12	9:gb_est13
10:gb_est14	11:gb_est15	12:gb_est16	13:gb_est17	14:gb_est18
15:gb_est19	16:gb_est2	17:gb_est20	18:gb_est21	19:gb_est22
20:gb_est23	21:gb_est24	22:gb_est25	23:gb_est26	24:gb_est27
25:gb_est28	26:gb_est29	27:gb_est30	28:gb_est31	29:gb_est32
30:gb_est33	31:gb_est34	32:gb_est35	33:gb_est36	34:gb_est37
35:gb_est38	36:gb_est39	37:gb_est40	38:gb_est41	39:gb_est42
40:gb_est43	41:gb_est44	42:gb_est45	43:gb_est46	44:gb_est47
45:gb_est48	46:gb_est49	47:gb_est50	48:gb_est51	49:gb_est52
50:gb_est53	51:gb_est54	52:gb_est55	53:gb_est56	54:gb_est57
55:gb_est58	56:gb_est59	57:gb_est60	58:gb_est61	59:gb_est62
60:gb_est63	61:gb_est64	62:gb_est65	63:gb_est66	64:gb_est67
65:gb_est68	66:gb_est69	67:gb_est70	68:gb_est71	69:gb_est72
70:gb_est73	71:gb_est74	72:gb_est75	73:gb_est76	74:gb_est77
75:gb_est78	76:gb_est79	77:gb_est80	78:gb_est81	79:gb_est82
80:gb_est83	81:gb_est84	82:gb_est85	83:gb_est86	84:gb_est87
85:gb_est88	86:gb_est89	87:gb_est90	88:gb_est91	89:gb_est92
90:gb_est93	91:gb_est94	92:gb_est95	93:gb_est96	94:gb_est97
95:gb_est98	96:gb_est99	97:gb_est100	98:gb_est101	99:gb_est102
100:gb_est103	101:gb_est104	102:gb_est105	103:gb_est106	104:gb_est107
105:gb_est108	106:gb_est109	107:gb_est110	108:gb_est111	109:gb_est112
110:gb_est113	111:gb_est114	112:gb_est115	113:gb_est116	114:gb_est117
115:gb_est118	116:gb_est119	117:gb_est120	118:gb_est121	119:gb_est122
120:gb_est123	121:gb_est124	122:gb_est125	123:gb_est126	124:gb_est127
125:gb_est128	126:gb_est129	127:gb_est130	128:gb_est131	129:gb_est132
130:gb_est133	131:gb_est134	132:gb_est135	133:gb_est136	134:gb_est137
135:gb_est138	136:gb_est139	137:gb_est140	138:gb_est141	139:gb_est142
140:gb_est143	141:gb_est144	142:gb_est145	143:gb_est146	144:gb_est147
145:gb_est148	146:gb_est149	147:gb_est150	148:gb_est151	149:gb_est152
150:gb_est153	151:gb_est154	152:gb_est155	153:gb_est156	154:gb_est157
155:gb_est158	156:gb_est159	157:gb_est160	158:gb_est161	159:gb_est162
160:gb_est163	161:gb_est164	162:gb_est165	163:gb_est166	164:gb_est167
165:gb_est168	166:gb_est169	167:gb_est170	168:gb_est171	169:gb_est172
170:gb_est173	171:gb_est174	172:gb_est175	173:gb_est176	174:gb_est177
175:gb_est178	176:gb_est179	177:gb_est180	178:gb_est181	179:gb_est182
180:gb_est183	181:gb_est184	182:gb_est185	183:gb_est186	184:gb_est187
185:gb_est188	186:gb_est189	187:gb_est190	188:gb_est191	189:gb_est192
190:gb_est193	191:gb_est194	192:gb_est195	193:gb_est196	194:gb_est197
195:gb_est198	196:gb_est199	197:gb_est200	198:gb_est201	199:gb_est202
200:gb_est203	201:gb_est204	202:gb_est205	203:gb_est206	204:gb_est207
205:gb_est208	206:gb_est209	207:gb_est210	208:gb_est211	209:gb_est212
210:gb_est213	211:gb_est214	212:gb_est215	213:gb_est216	214:gb_est217
215:gb_est218	216:gb_est219	217:gb_est220	218:gb_est221	219:gb_est222
220:gb_est223	221:gb_est224	222:gb_est225	223:gb_est226	224:gb_est227
225:gb_est228	226:gb_est229	227:gb_est230	228:gb_est231	229:gb_est232
230:gb_est233	231:gb_est234	232:gb_est235	233:gb_est236	234:gb_est237
235:gb_est238	236:gb_est239	237:gb_est240	238:gb_est241	239:gb_est242
240:gb_est243	241:gb_est244	242:gb_est245	243:gb_est246	244:gb_est247
245:gb_est248	246:gb_est249	247:gb_est250	248:gb_est251	249:gb_est252
250:gb_est253	251:gb_est254	252:gb_est255	253:gb_est256	254:gb_est257
255:gb_est258	256:gb_est259	257:gb_est260	258:gb_est261	259:gb_est262
260:gb_est263	261:gb_est264	262:gb_est265	263:gb_est266	264:gb_est267
265:gb_est268	266:gb_est269	267:gb_est270	268:gb_est271	269:gb_est272
270:gb_est273	271:gb_est274	272:gb_est275	273:gb_est276	274:gb_est277
275:gb_est278	276:gb_est279	277:gb_est280	278:gb_est281	279:gb_est282
280:gb_est283	281:gb_est284	282:gb_est285	283:gb_est286	284:gb_est287
285:gb_est288	286:gb_est289	287:gb_est290	288:gb_est291	289:gb_est292
290:gb_est293	291:gb_est294	292:gb_est295	293:gb_est296	294:

Statistics: Mean 12.051: Variance 1.914: scale 6.295

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARY

Result No.	Query			Length	DR	ID	Description	Pred. No.
	Score	Match	%					
1	564	19.7	711	26	AS004508	Homo sapiens	genomic D	0.00e+00
2	564	19.7	711	27	AG010947	Homo sapiens	genomic D	0.00e+00
3	567	19.5	721	27	AG011001	Homo sapiens	genomic D	0.00e+00
4	635	19.3	721	25	A0345532	Homo sapiens	genomic D	0.00e+00
5	635	18.9	718	26	AS004363	Homo sapiens	genomic D	0.00e+00
6	635	18.9	718	27	AG010489	Homo sapiens	genomic D	0.00e+00
7	530	18.7	648	26	AG004787	Homo sapiens	genomic D	0.00e+00
8	530	18.7	698	27	AS000976	Homo sapiens	genomic D	0.00e+00
9	634	18.6	696	27	AG009765	Homo sapiens	genomic D	0.00e+00
10	624	18.6	696	26	AG003576	Homo sapiens	genomic D	0.00e+00
11	564	19.6	720	28	AG013858	Homo sapiens	genomic D	0.00e+00
12	564	19.6	720	24	AG006462	Homo sapiens	genomic D	0.00e+00
13	519	18.4	747	26	AG007052	Homo sapiens	genomic D	0.00e+00



```

REFERENCE
AUTHORS      Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE        Homo sapiens genomic DNA, chromosome 21q
JOURNAL      Published Only in DataBase (1998) In press
REFERENCE
AUTHORS      Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE        Direct Submission
JOURNAL      Submitted (29-JUL-1998) to the DDBJ/EMBL/GenBank databases.
AUTHORS      Masakiya Harfuri, Kitasato University, Department of Science, JST
TITLE        Masakiya Harfuri, Kitasato University, Department of Science, JST
JOURNAL      Sequencing Laboratory: Kitasato U-15-1, Sagamihara 228, Japan
AUTHORS      (E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
TITLE        Fax:0427-78-9561)
FEATURES
            Location/Qualifiers
                1..721
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="21"
                /clone="879c12x91"
                /map="21q"
BASE COUNT      195 a 193 c 175 g 179 t 9 others
ORIGIN
Query Match      19.5%: Score 657: DB 27: Length 721:
Best local Similarity 98.7%: Pred. No. 0.00e+00:
Matches 983: Conservative 0: Mismatches 5: Indels 4: Gaps 4:

Db 32 CATCTTACGGATGGCATGACAGTAAGAGAAATTATCAGCTGCTGCCATAACCATGAGTGTAT 91
Cp 1772 CATCTTACGGATGGCATGACAGTAAGAGAAATTATCAGCTGCTGCCATAACCATGAGTGTAT 1713
Db 92 AACACTCGCGCAACTTACTTCTGCACACGATCGSAGCACCAGGAGCTTAACCGCTTTT 151
Cp 1712 AACACTCGCGCAACTTACTTCTGCACACGATCGGAGGACCGAGGAGCTTAACCGCTTTT 1653
Db 152 TTGCATCAATAGGGGATCATGATCTGCTTCTGCACACGATCGGAGGACCGAGGAGCTTAACCGCTTTT 211
Cp 1652 TTGCATCAATAGGGGATCATGATCTGCTTCTGCACACGATCGGAGGACCGAGGAGCTTAACCGCTTTT 1593
Db 212 GGCATACCAACGACGAGCGTGCACACGAGTGCCTGTAGCAATGGCAACACGCTTGCGC 271
Cp 1592 GGCATACCAACGAGTGCACACGAGTGCCTGTAGCAATGGCAACACGCTTGCGC 1533
Db 272 AAATCTATTAACTGGCGAACTACTTACTTACGCTTCCCGGCAACGATTAATAGACTGGATG 331
Cp 1532 AAATCTATTAACTGGCGAACTACTTACTTACGCTTCCCGGCAACGATTAATAGACTGGATG 1473
Db 332 GATGCTGATAAAGTTCAGAGACCACTTCTGCGCTGGCGCTTCGCGCTGGCTGTTTATT 391
Cp 1472 GATGCTGATAAAGTTCAGAGACCACTTCTGCGCTGGCGCTTCGCGCTGGCTGTTTATT 1413
Db 392 GCTGATAAATCTGACACGGTGGAGCTGGGCTTCGCGCTATCATTCGACGACTGGGGCCA 451
Cp 1412 GCTGATAAATCTGACACGGTGGAGCTGGGCTTCGCGCTATCATTCGACGACTGGGGCCA 1353
Db 452 GATGCTGATAAAGTTCAGAGACCACTTCTGCGCTGGCGCTTCGCGCTGGCTGTTTATT 511
Cp 1352 GATGCTGATAAAGTTCAGAGACCACTTCTGCGCTGGCGCTTCGCGCTGGCTGTTTATT 1293
Db 512 GAACTGATGACAGATGCTGATAGTGGCTCTACTGATTAAGC-TTGCTAACTGTCA 570
Cp 1292 GAACTGATGACAGATGCTGATAGTGGCTCTACTGATTAAGC-TTGCTAACTGTCA 1233
Db 571 GACCAAGTTTACTC-TATATATCTTTAGATTCGATTTAAACCTCATTTTTTAATTTAAAGG 629
Cp 1232 GACCAAGTTTACTC-TATATATCTTTAGATTCGATTTAAACCTCATTTTTTAATTTAAAGG 1173
Db 629 ATCTAGGTGAAGATCCTTTT-GATAATCATGACCAA-TCCCTTAACGTTGAGTTTCG 687
Cp 1172 ATCTAGGTGAAGATCCTTTT-GATAATCATGACCAA-TCCCTTAACGTTGAGTTTCG 1113
Db 688 TTCCACTGAGCTCAGANCCGTTAGAAAAA 719
Cp 1112 TTCCACTGAGCTCAGANCCGTTAGAAAAA 1081

```

RESULT	4	AC004662	721 bp	DNA	GSS	30-JAN-1998
LOCUS						
DEFINITION						Homo sapiens genomic DNA, 21q region, clone: g79G12X91, genomic survey sequence.
ACCESSION		AG004662				
NID		g2826191				
KEYWORDS		GSS.				
ORGANISM		Homo sapiens				
SOURCE		Homo sapiens				
		Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;				
		Primates; Catarrhini; Hominoidea; Hominidae; Homo.				
REFERENCE		1 (bases 1 to 721)				
AUTHORS		Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.				
TITLE		Homo sapiens genomic DNA, chromosome 21q				
JOURNAL		Published Only in Database (1998), In press				
REFERENCE		2 (bases 1 to 721)				
AUTHORS		Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.				
TITLE		Direct Submmission				
JOURNAL		Submitted (30-JAN-1998) to the DBJ/EMBL/GenBank databases.				
		Masahira Hattori, Kitasato University, Department of Science, JST Sequencing Laboratory, Kitasato i-15-1, Sagamihara 228, Japan				
		(E-mail:hattori@hgcm.sims.u-tokyo.ac.jp, Tel:0427-78-9732, Fax:0427-78-9561)				
FEATURES		Location/Qualifiers				
source		1..721				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
		/chromosome="21"				
		/clone="g79G12X91"				
		/map="21q"				
BASE COUNT		195 a 163 c 175 g 179 t 9 others				
ORIGIN						
Query Match		19.5%; Score 657; DB 26; Length 721;				
Best Local Similarity		98.7%, Pred No 0.00e+00;				
Matches		683; Conservative 0; Mismatches 5; Indels 4; Gaps 4				
Db	32	CATCTTAGCGGATGCATGACGTAGAGAAATTATCCAGTCTGCCATAACCATGAGTGAT	91			
Cp	1772	CATCTTAGCGGATGCATGACGTAGAGAAATTATTCAGTCTGCCATAACCATGAGTGAT	1713			
Db	92	AACACTGGGGCACTTACTTCTGACAACGATCGGAGAACGAAGAGAGTAACTGGCTTT	151			
Cp	1712	AACACTGGGGCACTTACTTCTGACAACGATCGGAGAACGAAGAGAGTAACTGGCTTT	1653			
Db	152	TTGCACAAATGGGGATCATGTTAAGTCTGCAACGATCGGAGAACGGAGCTGAATGAA	211			
Cp	1652	TTGCACAAATGGGGATCATGTTAAGTCTGCAACGATCGGAGAACGGAGCTGAATGAA	1593			
Db	212	GCCATACCAAAGCAGAGCGTGACACCAAGATGCTGTAGCAATGGCAAAAGAAGTGGG	271			
Cp	1592	GCCATACCAAAGCAGAGCGTGACACCAAGATGCTGTAGCAATGGCAAAAGAAGTGGG	1533			
Db	272	AAACTATTACTGGCGAAGTACTTACTTCTAGCTTCGGGCAACAAATTAATAGCTGGATG	331			
Cp	1532	AAACTATTACTGGCGAAGTACTTACTTCTAGCTTCGGGCAACAAATTAATAGCTGGATG	1473			
Db	332	GAGCGGATAAAGTTGACAGACCACTTCTGGCTTCGGCCCTTCGGCTGGCTTGGTTTATT	391			
Cp	1472	GAGCGGATAAAGTTGACAGACCACTTCTGGCTTCGGCCCTTCGGCTGGCTTGGTTTATT	1413			
Db	392	GCTGATAATCTGACGCGGTGAGCGGGGTCTGGCGATCATTCATTCAGATCGAGGCA	451			
Cp	1412	GCTGATAATCTGACGCGGTGAGCGGGGTCTGGCGATCATTCATTCAGATCGAGGCA	1353			
Db	452	GATGGTAAAGCCCTCCGCTATCGTAGTTATCTACAGGACGGGAGTCAGGCAACTATGAT	511			
Cp	1352	GATGGTAAAGCCCTCCGCTATCGTAGTTATCTACAGGACGGGAGTCAGGCAACTATGAT	1293			
Db	512	GACGAAAATAGACAGATCGCTGAGATAGGTGGCTCACTGATTAGCCTTGGTAACCTGCA	570			







Cp 1442 GCGTGGGCGCTT-CGCGCTGTC-ATGCTTTATGCTGTAATAATGTAAGAGCGCGGTGAGCGTG 1385

Db 590 GTCTC 595  
|||||

Cp 1384 GTCTC 1379

RESULT 10  
LOCUS AG03576 596 bp DNA GSS 19-DEC-1997  
DEFINITION Homo sapiens genomic DNA, 21q region, clone: p8G4SPN15, genomic survey sequence.

ACCESSION AG03576  
NID g2706702  
KEYWORDS GSS.  
SOURCE Homo sapiens DNA, clone: p8G4SPN15.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 596)  
Hattori.M., Ishii.K., Shiba.T. and Sakaki.Y.  
TITLE Homo sapiens genomic DNA, chromosome 21q  
JOURNAL Published Only in Database (1997) In press  
REFERENCE 2 (bases 1 to 596)  
Hattori.M., Ishii.K., Shiba.T. and Sakaki.Y.  
AUTHORS  
TITLE Direct Submission  
JOURNAL Submitted (19-DEC-1997) to the DBJ/EMBL/GenBank databases.  
Masahira Hattori, Kitasato University, Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan (E-mail:hattori@hgc.ims.u-tokyo.ac.jp. Tel:0427-78-9732, Fax:0427-78-9561)

FEATURES  
source  
1..596  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/clone="p8G4SPN15"  
/map="21q"

BASE COUNT 183 a 165 c 185 g 162 t 1 others  
ORIGIN

Query Match 19.6%; Score 624; DB 26; Length 596.  
Best Local Similarity 99.1%; Pred No. 0.00e+00  
Matches 560; Conservative 0; Mismatches 0; Indels 6; Gaps 6;

Db 31 GTTTTCTCAGCCAGAAACGCTGCTGAAAGTAAGATGATGTAAGATCAGTTGGTGCA 90  
|||||

Cp 2039 GTTTTCTCAGCCAGAAACGCTGCTGAAAGTAAGATGATGTAAGATCAGTTGGTGCA 1980  
|||||

Db 91 CGACTGGTATACATCAATGTAAGTGAACACACACAGCGGTAGATCCTTGACAGTTTCGCCCC 150  
|||||

Cp 1979 CGAGTGGGTATCATGCAATGATGATCTCAACAGCGGTAGATCCTTGACAGTTTCGCCCC 1920  
|||||

Db 151 GAAGAAAGTTTTCAGTAAGTGAACACACACACACACACACACACACACACACACACACAC 210  
|||||

Cp 1919 GAAGAAAGTTTTCAGTAAGTGAACACACACACACACACACACACACACACACACACAC 1860  
|||||

Db 211 CGTATTGACGGCGGCAAGACAGCACTGCTGCGGCATACACTATTCTCAGAAATCAGTTG 270  
|||||

Cp 1859 GTATTGACGGCGGCAAGACAGCACTGCTGCGGCATACACTATTCTCAGAAATCAGTTG 1800  
|||||

Db 271 GTTCACTACTCAGTACACAGAAAGCATCTTACGGATGGGCATCAGCTAGAGAAATTA 330  
|||||

Cp 1799 GTTCACTACTCAGTACACAGAAAGCATCTTACGGATGGGCATCAGCTAGAGAAATTA 1740  
|||||

Db 331 TCGAGTGTGCTGCCATACACATGATGATACACTGCGGCAACTTACTTCTGACACAGCATC 390  
|||||

Cp 1739 TCGAGTGTGCTGCCATACACATGATGATACACTGCGGCAACTTACTTCTGACACAGCATC 1680  
|||||

Db 391 GGAAGACCAAGACAGAAATTAATGCTTTTITTTTGTAAACATCGGGGATCATGTAACCTCGCCTT 450  
|||||

Cp 1679 GGAAGACCAAGACAGAAATTAATGCTTTTITTTTGTAAACATCGGGGATCATGTAACCTCGCCTT 1620  
|||||

Db 451 GATCGTTGGTAAATGCTGATGTAAGTGAACACACACACACACACACACACACACACACAC 510  
|||||

Cp 1619 GATCGTTGGAAACGTAAGTGAAGTGAACAGAGTGAACAGAGTGAACAGAGTGAACAGAGT 1560  
|||||

Db 511 GGCTGTAGCAATAGGCAATGAAGTTTGGGCAAAATATTAATTAATTAATTAATTAATTAATTA 570  
|||||

Cp 1560 G-CCTGTAGCAATAGGCAACAGTTTGGGCAAAATATTAATTAATTAATTAATTAATTAATTA 1520  
|||||

Db 571 CTTCGGGGGCAACAATTAATAGACTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 630  
|||||

Cp 1501 CTTCGGGGGCAACAATTAATAGACTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440  
|||||

Db 631 CGCTCGGCGCTTTTCGGGCTGGGCTGGGCTTTATTCGTAAGTGAAGTGAAGTGAAGTGAAGT 689  
|||||

Cp 1442 CGCTCGGCGCTTTTCGGGCTGGGCTGGGCTTTATTCGTAAGTGAAGTGAAGTGAAGTGAAGT 1385  
|||||

Db 590 GTCTC 595  
|||||

Cp 1384 GTCTC 1379

RESULT 11  
LOCUS AG013858 720 bp DNA GSS 10-SEP-1998  
DEFINITION Homo sapiens genomic DNA, 21q region, clone: 762015SpN14, genomic survey sequence.

ACCESSION AG013858  
NID g3560348  
KEYWORDS GSS.  
SOURCE Homo sapiens DNA, clone: 762015SpN14.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 720)  
Hattori.M., Ishii.K., Toyoda.A., Shiba.T. and Sakaki.Y.  
TITLE Homo sapiens genomic DNA, chromosome 21q  
JOURNAL Published Only in Database (1998) In press  
REFERENCE 2 (bases 1 to 720)  
Hattori.M., Ishii.K., Toyoda.A., Shiba.T. and Sakaki.Y.  
AUTHORS  
TITLE Direct Submission  
JOURNAL Submitted (10-SEP-1998) to the DBJ/EMBL/GenBank databases.  
Masahira Hattori, Kitasato University, Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan (E-mail:hattori@hgc.ims.u-tokyo.ac.jp. Tel:0427-78-9732, Fax:0427-78-9561)

FEATURES  
Location/Qualifiers  
source  
1..720  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/clone="762015SpN14"  
/map="21q"

BASE COUNT 172 a 178 c 170 g 186 t 14 others  
ORIGIN

Query Match 18.6%; Score 624; DB 28; Length 720.  
Best Local Similarity 97.2%; Pred No. 0.00e+00;  
Matches 671; Conservative 0; Mismatches 13; Indels 6; Gaps 6;

Db 35 AGCTCCCTCGTGGCGCTCTGCTGTTGCGACCTGTCGCTTACCGGATACCTGTCGCGGCTTT 94  
|||||

QY 586 AGCTCCCTCGTGGCGCTCTGCTGTTGCGACCTGTCGCTTACCGGATACCTGTCGCGGCTTT 645  
|||||

Db 95 CTCCTCTCGGSAAGCGTGGGGGCTTTTCTCATATCTATGATGATGATGATGATGATGATGATG 154  
|||||

QY 646 CTCCTCTCGGSAAGCGTGGGGGCTTTTCTCATATCTATGATGATGATGATGATGATGATGATG 705  
|||||

Db 155 TAGTCTCTGCTGCCACAGCTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 214  
|||||

QY 706 TAGTCTCTGCTGCCACAGCTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 765  
|||||

Db 215 GCCTTATCCGGTAACATGCTGCTTGTGATGCTGCAAGCGGTAAGACAGCACTTATCGGCACTG 274  
|||||

QY 766 GCCTTATCCGGTAACATGCTGCTTGTGATGCTGCAAGCGGTAAGACAGCACTTATCGGCACTG 345  
|||||









```
*****  
W P E L H  
  
Release 3.1A Jock F. Collins, BioComputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution Rights by Oxford Molecular Ltd  
(TM)  
*****
```

psrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm  
Run on: Sun Mar 14 10:56:35 1999; Maspar time 5168 90 Seconds  
1538.452 Million cell updates/sec.

Molecular output not generated.

AUS-09-020-716-2  
Description: G13365|LSDP USG0029716 seq  
Perfect Score: 3365  
A. Sequence: TCGACCTGGAGAAATGAAGAATA  
Comp: AGTGGAGACTGCACCCTGTTC  
Scoring table: TABLE default  
Gap: 5

Cloni	77.6	4824	32	AS14304
Cloni	76.0 <td>5350 <td>32 <td>AB003468</td> </td></td>	5350 <td>32 <td>AB003468</td> </td>	32 <td>AB003468</td>	AB003468
Cloni	75.3 <td>2850 <td>32 <td>SYNPT218U</td> </td></td>	2850 <td>32 <td>SYNPT218U</td> </td>	32 <td>SYNPT218U</td>	SYNPT218U

```

Searched: 502357 seqs. 1181590623 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries

emb157
1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
7:em_om 8:em_og 9:em_ov 10:em_pat 11:em_ph 12:em_pl
13:em_ro 14:em_un 15:em_v1
genbank110
16:qb_bal 17:qb_ba2 18:qb_htg 19:qb_in 20:qb_om 21:qb_ov
22:qb_pat 23:qb_ph 24:qb_pl 25:qb_pl2 26:qb_pr1
27:qb_pr2 28:qb_pr3 29:qb_ro 30:qb_st 31:qb_sts 32:qb_sy
33:qb_un 34:qb_v1

Statistics: Mean 11.932; Variance 5.489; scale 2.174

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result  % Query Length DB ID Description Pred. No.
No. Score Match
C 1 2907 85.4 4412 32 EVCOR112N Expression vector pCOR 0.00e+00
C 2 2897 85.1 5641 32 AF016404 Chloroplast transform 0.00e+00
C 3 2883 85.7 4433 32 EVCOR114N Expression vector pCOR 0.00e+00
C 4 2878 85.5 2967 32 G02449 Cloning vector pDIRECT 0.00e+00
C 5 2864 85.1 4133 32 G01668 Phagemid cloning vecto 0.00e+00
C 6 2786 82.8 6824 25 A25909 Yeast integration vect 0.00e+00
C 7 2737 81.3 5160 16 ECTTDN3 E.coli pT7GH pl DNA, 0.00e+00
C 8 2737 81.3 8953 16 ECTTDN2 E.coli pT7T pl DNA, 0.00e+00
C 9 2737 81.3 11271 16 ECTTDN1 E.coli (HM174) pT717/T 0.00e+00
C 10 2665 79.2 2863 32 AF013597 Phagemid cloning vecto 0.00e+00
C 11 2632 77.6 4864 32 ASY14584 Artificial sequence; v 0.00e+00
C 12 2557 76.0 5350 32 AB003468 Cloning vector pAP3neo 0.00e+00

```

































[illegible]



















[illegible]



[illegible]



Cp 1760 ACAGATGGGATGACACTAAGCAATTAATGACATGCTGTCATTAACCATGATGATACACT 1707  
 Db 1456 GGGGCTAAATTAATTAACAGCATGCGAAGAGAGAGCTAAACGGCTTTTTCAC 1515  
 Cp 1766 GGGGCTAAATTAATTAACAGCATGCGAAGAGAGAGCTAAACGGCTTTTTCAC 1647  
 Db 1516 AATAGGAGGATCATGATTAACCTGGCTTATGATGAGGAGGAGCTGATGAGGAGG 1575  
 Cp 1646 AATAGGAGGATCATGATTAACCTGGCTTATGATGAGGAGGAGCTGATGAGGAGG 1587  
 Db 1576 GGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1635  
 Cp 1586 GGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1527  
 Db 1636 TTAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1695  
 Cp 1526 TTAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1467  
 Db 1696 GATAAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1755  
 Cp 1466 GATAAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1407  
 Db 1756 AATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1815  
 Cp 1406 AATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1347  
 Db 1816 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1875  
 Cp 1346 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1287  
 Db 1876 AATAGAGAGATGCGTGGAGATGCGTGGAGATGCGTGGAGATGCGTGGAGATGCGTGGAG 1935  
 Cp 1286 AATAGAGAGATGCGTGGAGATGCGTGGAGATGCGTGGAGATGCGTGGAGATGCGTGGAG 1227  
 Db 1936 GTTACGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1995  
 Cp 1226 GTTACGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1167  
 Db 1995 GTTACGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2055  
 Cp 1166 GTTACGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1107  
 Db 2055 TGAGGAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2115  
 Cp 1106 IGAGGAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1047  
 Db 2115 GTTACGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2175  
 Cp 1046 GTTACGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 987  
 Db 2175 CAAGAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2235  
 Cp 986 CAAGAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 927  
 Db 2235 AGTGGCTTCTAGTGTAGGCGGTAGTGTAGGCGGTAGTGTAGGCGGTAGTGTAGGCGGTAGTGT 2295  
 Cp 926 AGTGGCTTCTAGTGTAGGCGGTAGTGTAGGCGGTAGTGTAGGCGGTAGTGTAGGCGGTAGTGT 867  
 Db 2295 ACAATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2355  
 Cp 866 ACAATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 807  
 Db 2356 GTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2415  
 Cp 806 GTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 747  
 Db 2415 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2475  
 Cp 746 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 687  
 Db 2476 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2535  
 Cp 586 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 627

Db 2535 GTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2595  
 Cp 526 GTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 567  
 Db 2596 TATCTTTATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2655  
 Cp 566 TATCTTTATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 507  
 Db 2656 TCGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2715  
 Cp 506 TCGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 447  
 Db 2716 GCTTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2775  
 Cp 446 GCTTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 387  
 Db 2776 AACCGTATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2835  
 Cp 386 AACCGTATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 327  
 Db 2836 AGCGAGTCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2895  
 Cp 326 AGCGAGTCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 327

RESULT 15  
 ID E11270 standard: DNA, UNC: 2927 BP.  
 AC E11270:  
 NI d1109607  
 DT 08-OCT-1997 (Rel. 52, Created)  
 DT 08-OCT-1997 (Rel. 52, last updated, Version 1)  
 DE Nucleotide sequence of pTRP.  
 KW JP 199610379-A/2  
 OS unidentified  
 OS unclassified.  
 RN [1]  
 RP 1-2927  
 RA Ishizuka T., Kihira Y., Tanaka T., Matsuo T.,  
 "PRODUCTION OF RECOMBINED HUMAN MYOGLOBIN".  
 RL Patent number JP 199610379-A/2, 23-APR-1996.  
 PL ORIENTAL YEAST CO LTD.  
 OS None  
 CC OC Artificial sequences.  
 CC OC JP 199610379-A/2  
 CC PD 23-APR-1996  
 CC PF 07-OCT-1994 JP 1994268121  
 CC PI ISHIZUKA TOSHIO, KIHIRA YASUNORI, TANAKA TOSHIO,  
 MATSUO TAKESHI  
 CC PC G12N15/09, C07H21/04, C07K14/805, G12P21/02, G12P21/02, G12P21/19)  
 CC CC strandedness: Double;  
 CC CC topology: Circular;  
 CC CC key Location/Qualifiers  
 CC FT source 1..2927  
 CC FT /organism="Artificial sequences"  
 CC FT promoter 267..326  
 CC FT /note="trp promoter derived from E. coli K12"  
 CC FT CDS 1127..1197  
 CC FT /product="beta-lactamase"  
 CC FT key Location/Qualifiers  
 FH FT source 1..2927  
 FT /organism="unidentified"  
 FT Sequence 2927 BP: 735 A: 736 G: 718 T: 0 other:  
 SQ

Query Match 75.1%, Score 2529, DB 10: Length 2927.  
 Best Local Similarity 59.7%, E-Val 0.00e+00;  
 Matches 2541: Conservative 0; Mismatches 6; Indels 1; Gaps 1;  
 Db 380 TCAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 439  
 Cp 2845 TCAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2786



US-09-020-716-2.rge

Wed Mar 17 09:43:27 1999

```

Cp 526 3TAACTCCACGCTCGGAAAGAGAGAGCTTACAGAGAGCTTCCAGGGGAAACCCCTGG 567
Db 2680 TATCTTATATAGTCTGTGGGTTTCGCCAGCTCTGACTTGAAGCGTCGATTTTGTGATGC 2719
Cp 566 TATCTTATATAGTCTGTGGGTTTCGCCAGCTCTGACTTGAAGCGTCGATTTTGTGATGC 507
Db 2720 TCGTCAGGGGGGCGGAGCCTATGGAAGAGCGGAGCAAGCGGCTTTTACGGTTCCCTG 2779
Cp 506 TCGTCAGGGGGGCGGAGCCTATGGAAGAGCGGAGCAAGCGGCTTTTACGGTTCCCTG 447
Db 2780 GGGTTTGGTGGGCTTTTGGCTCAGATGTTCTTCTCGGTTATCCGCTGATTTCTGTGGAT 2839
Cp 446 GGGTTTGGTGGGCTTTTGGCTCAGATGTTCTTCTCGGTTATCCGCTGATTTCTGTGGAT 387
Db 2840 AACGCTATTACGGCTTTTGGCTCAGATGTTCTTCTCGGTTATCCGCTGATTTCTGTGGAT 2899
Cp 386 AACGCTATTACGGCTTTTGGCTCAGATGTTCTTCTCGGTTATCCGCTGATTTCTGTGGAT 327
Db 2900 AACGCTATTACGGCTTTTGGCTCAGATGTTCTTCTCGGTTATCCGCTGATTTCTGTGGAT 2927
Cp 326 AACGCTATTACGGCTTTTGGCTCAGATGTTCTTCTCGGTTATCCGCTGATTTCTGTGGAT 299

```

Search completed: Sun Mar 14 13:14:36 1999  
Job time : 8281 secs.





\*\*\*\*\*

WIRELH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution Rights by Oxford Molecular Ltd

MPsrch\_n 0 1 - n a database search, using Smith-Waterman algorithm  
Run on: Sun Mar 14 15:24:16 1999: Maspar time 485.47 Seconds  
943.032 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-020-716-2  
Description: (1:3365) from US09020716 seq  
Perfect Score: 3365  
N.A. Sequence: 1 TCACCTCGAGGGGGGGGGC  
Comp: AGCTGAGTTCGCGCGCGG

Scoring table:  
Gap 5  
TABLE default.

Nmatch STD : Dbase 0: Query 0  
Searched: 188442 seqs, 58026449 bases x 2  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-geneseq2  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39 40:part40

Statistics: Mean 10.017; Variance 5.937; scale 1.687

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2855	84.8	534	25	PUMIGIT sequence incl	0.00e+00
2	2844	83.3	728	14	Plasmid pWP3169 enco	0.00e+00
3	2820	83.8	558	14	Plasmid pXP6-IL2	0.00e+00
4	2818	83.7	558	29	Interleukin-2 express	0.00e+00
5	2774	82.4	5924	6	K lactis/S. cerevisae	0.00e+00
6	2773	75.2	292	24	Plasmid pTRP	0.00e+00
7	2772	75.2	292	24	Plasmid pTRP	0.00e+00
8	2752	71.2	524	34	Vector pET Trc S05/N	0.00e+00
9	2745	71.2	524	17	Plasmid pET Trc S05/N	0.00e+00
10	2740	70.7	3554	11	Plasmid pASK46 for ex	0.00e+00
11	2735	70.5	3714	13	pW324b Plasmid expre	0.00e+00
12	2734	67.0	3812	11	Plasmid pASK60-Strep	0.00e+00
13	2754	67.0	4515	8	pLISC-SE vector.	0.00e+00

C	14	2254	67.0	4640	5	Q26910	RBI, PDI coexpression	0.00e+00
C	15	2222	65.0	5178	27	T49876	Pret-Splice	0.00e+00
C	16	2220	65.0	6205	27	T49877	Autoregulatory vector	0.00e+00
C	17	2222	66.0	7474	31	T71320	Plasmid pT73 encodin	0.00e+00
C	18	2215	65.8	6414	3	Q26733	PAD-CMV1 expression v	0.00e+00
C	19	2215	65.8	6414	1	Q36283	PAD-CMV1 contig a Tum	0.00e+00
C	20	2203	65.5	3681	2	Q13578	Plasmid pKSEL5	0.00e+00
C	21	2205	65.5	5314	2	Q13576	Plasmid pMTV1 contain	0.00e+00
C	22	2203	65.5	6414	3	Q20766	PAD-CMV1	0.00e+00
C	23	2202	65.4	3699	39	V14340	Plasmid pASGFP expres	0.00e+00
C	24	2202	65.4	4145	7	Q40279	Sequence of clone pS2	0.00e+00
C	25	2200	65.4	4792	14	Q87347	Plasmid pINVI	0.00e+00
C	26	2202	65.4	4792	15	Q84696	Plasmid GS contg. N.	0.00e+00
C	27	2202	65.4	5042	15	Q84694	Plasmid glucamylase	0.00e+00
C	28	2202	65.4	5356	25	T43794	Plasmid pPRIPAT (rat	0.00e+00
C	29	2200	65.4	10930	20	Q81226	Plasmid pM16-1	0.00e+00
C	30	2200	65.4	10950	20	Q81225	Plasmid pM16	0.00e+00
C	31	2198	65.3	4277	7	Q40280	Sequence of clone pS2	0.00e+00
C	32	2198	65.3	6926	7	Q40281	Sequence of clone pS2	0.00e+00
C	33	2195	65.2	14311	24	T38744	Plasmid pSVFL(-) inc	0.00e+00
C	34	2187	65.0	7238	9	Q50201	Vector p37M1-10C	0.00e+00
C	35	2170	64.5	4118	35	T69189	Construct pGEM-nr(Ka	0.00e+00
C	36	2170	64.5	4118	35	T69188	Construct pGEM-nr(Ka	0.00e+00
C	37	2171	64.5	6295	39	V02043	Plasmid pWRG3195 enco	0.00e+00
C	38	2166	64.4	4540	32	T69891	Human cyclin D1-human	0.00e+00
C	39	2166	64.4	4824	20	T08975	Shortened C1 gene in	0.00e+00
C	40	2167	64.4	4883	4	Q25709	pDEL10	0.00e+00
C	41	2167	64.4	4883	8	Q51192	pDEL10; plasmid DNA i	0.00e+00
C	42	2167	64.4	5620	4	Q27489	Plasmid pVE108 used i	0.00e+00
C	43	2167	64.4	5620	4	Q27489	pVE108	0.00e+00
C	44	2167	64.4	5642	8	Q51193	pDE4; plasmid DNA rep	0.00e+00
C	45	2167	64.4	7050	7	Q40419	Sequence of pTE4.	0.00e+00

ALIGNMENTS

RESULT 1  
T43137 standard: DNA: 5534 BP.

AC T43137:  
DE 10-FEB-1997 (first entry)  
DE PUMIGIT sequence including upstream activating sequence.  
KW Transgenic plant, gene expression, upstream activating sequence;  
KW UAS: transactivating protein, Gal4; herbicide resistance,  
KW polyhydroxybutyrate; safety, pUMIGIT; ds.  
OS Synthetic.  
PN CA2150039-A.  
PD 09-AUG-1996.  
PF 24-MAY-1995; 150039.  
PR 08-FEB-1995; GS-002456.  
PR (UYWA-) UNIV WARWICK.  
PA Bennett M, May S, Ramsay N;  
PI WPI: 96-486150/49.  
DR Control of genes in transgenic plants - using an upstream activating  
PT sequence activated by a transactivating protein expressed using a  
PT separate promoter  
PT  
PS Example; Fig 5F: 48pp: English.  
PS Reporter plasmid pUMIGIT (T43137), or pUAS Minimal promoter Gus  
CC Inot Terminator, contains a beta-glucuronidase (GUS) reporter  
CC gene under the control of the 46S cauliflower-mosaic virus minimal  
CC promoter and 10 synthetic 17-bp GAL4 binding sites. The upstream  
CC activating sequence (UAS) of pUMIGIT is activatable by yeast  
CC transactivating protein GAL4. In a novel method for controlling  
CC gene expression, a first transgenic plant carrying a gene encoding  
CC a desired phenotype (herbicide resistance of polyhydroxybutyrate  
CC prodn.) operatively linked to a UAS recognition site is pollinated  
CC by a second transgenic plant carrying Gal4 DNA (see also T43136).  
CC The transgene is fully expressed in F1 hybrid plants but segregates  
CC apart in subsequent generations, improving environmental safety.  
SQ Sequence 5534 BP. 1400 A. 1355 C. 1487 G. 1363 T.

Query Match 84.8%. Score 2955. FE 25. Length 5534;  
Best Local Similarity 100.0%. Pred. No. 0.00e+00;  
Matches 2855, Conservative 0, Mismatches 0, Indels 0, Gaps 0







































```

3496 3555 3615 3675 3735 3795 3855 3915 3975 4035 4095 4155 4215 4275 4335 4395 4455 4515 4575 4635 4695 4755 4815 4875 4935 4995 5055 5115 5175 5235 5295 5355 5415 5475 5535 5595 5655 5715 5775 5835 5895 5955 6015 6075 6135 6195 6255 6315 6375 6435 6495 6555 6615 6675 6735 6795 6855 6915 6975 7035 7095 7155 7215 7275 7335 7395 7455 7515 7575 7635 7695 7755 7815 7875 7935 7995 8055 8115 8175 8235 8295 8355 8415 8475 8535 8595 8655 8715 8775 8835 8895 8955 9015 9075 9135 9195 9255 9315 9375 9435 9495 9555 9615 9675 9735 9795 9855 9915 9975 10035 10095 10155 10215 10275 10335 10395 10455 10515 10575 10635 10695 10755 10815 10875 10935 10995 11055 11115 11175 11235 11295 11355 11415 11475 11535 11595 11655 11715 11775 11835 11895 11955 12015 12075 12135 12195 12255 12315 12375 12435 12495 12555 12615 12675 12735 12795 12855 12915 12975 13035 13095 13155 13215 13275 13335 13395 13455 13515 13575 13635 13695 13755 13815 13875 13935 13995 14055 14115 14175 14235 14295 14355 14415 14475 14535 14595 14655 14715 14775 14835 14895 14955 15015 15075 15135 15195 15255 15315 15375 15435 15495 15555 15615 15675 15735 15795 15855 15915 15975 16035 16095 16155 16215 16275 16335 16395 16455 16515 16575 16635 16695 16755 16815 16875 16935 16995 17055 17115 17175 17235 17295 17355 17415 17475 17535 17595 17655 17715 17775 17835 17895 17955 18015 18075 18135 18195 18255 18315 18375 18435 18495 18555 18615 18675 18735 18795 18855 18915 18975 19035 19095 19155 19215 19275 19335 19395 19455 19515 19575 19635 19695 19755 19815 19875 19935 19995 20055 20115 20175 20235 20295 20355 20415 20475 20535 20595 20655 20715 20775 20835 20895 20955 21015 21075 21135 21195 21255 21315 21375 21435 21495 21555 21615 21675 21735 21795 21855 21915 21975 22035 22095 22155 22215 22275 22335 22395 22455 22515 22575 22635 22695 22755 22815 22875 22935 22995 23055 23115 23175 23235 23295 23355 23415 23475 23535 23595 23655 23715 23775 23835 23895 23955 24015 24075 24135 24195 24255 24315 24375 24435 24495 24555 24615 24675 24735 24795 24855 24915 24975 25035 25095 25155 25215 25275 25335 25395 25455 25515 25575 25635 25695 25755 25815 25875 25935 25995 26055 26115 26175 26235 26295 26355 26415 26475 26535 26595 26655 26715 26775 26835 26895 26955 27015 27075 27135 27195 27255 27315 27375 27435 27495 27555 27615 27675 27735 27795 27855 27915 27975 28035 28095 28155 28215 28275 28335 28395 28455 28515 28575 28635 28695 28755 28815 28875 28935 28995 29055 29115 29175 29235 29295 29355 29415 29475 29535 29595 29655 29715 29775 29835 29895 29955 30015 30075 30135 30195 30255 30315 30375 30435 30495 30555 30615 30675 30735 30795 30855 30915 30975 31035 31095 31155 31215 31275 31335 31395 31455 31515 31575 31635 31695 31755 31815 31875 31935 31995 32055 32115 32175 32235 32295 32355 32415 32475 32535 32595 32655 32715 32775 32835 32895 32955 33015 33075 33135 33195 33255 33315 33375 33435 33495 33555 33615 33675 33735 33795 33855 33915 33975 34035 34095 34155 34215 34275 34335 34395 34455 34515 34575 34635 34695 34755 34815 34875 34935 34995 35055 35115 35175 35235 35295 35355 35415 35475 35535 35595 35655 35715 35775 35835 35895 35955 36015 36075 36135 36195 36255 36315 36375 36435 36495 36555 36615 36675 36735 36795 36855 36915 36975 37035 37095 37155 37215 37275 37335 37395 37455 37515 37575 37635 37695 37755 37815 37875 37935 37995 38055 38115 38175 38235 38295 38355 38415 38475 38535 38595 38655 38715 38775 38835 38895 38955 39015 39075 39135 39195 39255 39315 39375 39435 39495 39555 39615 39675 39735 39795 39855 39915 39975 40035 40095 40155 40215 40275 40335 40395 40455 40515 40575 40635 40695 40755 40815 40875 40935 40995 41055 41115 41175 41235 41295 41355 41415 41475 41535 41595 41655 41715 41775 41835 41895 41955 42015 42075 42135 42195 42255 42315 42375 42435 42495 42555 42615 42675 42735 42795 42855 42915 42975 43035 43095 43155 43215 43275 43335 43395 43455 43515 43575 43635 43695 43755 43815 43875 43935 43995 44055 44115 44175 44235 44295 44355 44415 44475 44535 44595 44655 44715 44775 44835 44895 44955 45015 45075 45135 45195 45255 45315 45375 45435 45495 45555 45615 45675 45735 45795 45855 45915 45975 46035 46095 46155 46215 46275 46335 46395 46455 46515 46575 46635 46695 46755 46815 46875 46935 46995 47055 47115 47175 47235 47295 47355 47415 47475 47535 47595 47655 47715 47775 47835 47895 47955 48015 48075 48135 48195 48255 48315 48375 48435 48495 48555 48615 48675 48735 48795 48855 48915 48975 49035 49095 49155 49215 49275 49335 49395 49455 49515 49575 49635 49695 49755 49815 49875 49935 49995 50055 50115 50175 50235 50295 50355 50415 50475 50535 50595 50655 50715 50775 50835 50895 50955 51015 51075 51135 51195 51255 51315 51375 51435 51495 51555 51615 51675 51735 51795 51855 51915 51975 52035 52095 52155 52215 52275 52335 52395 52455 52515 52575 52635 52695 52755 52815 52875 52935 52995 53055 53115 53175 53235 53295 53355 53415 53475 53535 53595 53655 53715 53775 53835 53895 53955 54015 54075 54135 54195 54255 54315 54375 54435 54495 54555 54615 54675 54735 54795 54855 54915 54975 55035 55095 55155 55215 55275 55335 55395 55455 55515 55575 55635 55695 55755 55815 55875 55935 55995 56055 56115 56175 56235 56295 56355 56415 56475 56535 56595 56655 56715 56775 56835 56895 56955 57015 57075 57135 57195 57255 57315 57375 57435 57495 57555 57615 57675 57735 57795 57855 57915 57975 58035 58095 58155 58215 58275 58335 58395 58455 58515 58575 58635 58695 58755 58815 58875 58935 58995 59055 59115 59175 59235 59295 59355 59415 59475 59535 59595 59655 59715 59775 59835 59895 59955 60015 60075 60135 60195 60255 60315 60375 60435 60495 60555 60615 60675 60735 60795 60855 60915 60975 61035 61095 61155 61215 61275 61335 61395 61455 61515 61575 61635 61695 61755 61815 61875 61935 61995 62055 62115 62175 62235 62295 62355 62415 62475 62535 62595 62655 62715 62775 62835 62895 62955 63015 63075 63135 63195 63255 63315 63375 63435 63495 63555 63615 63675 63735 63795 63855 63915 63975 64035 64095 64155 64215 64275 64335 64395 64455 64515 64575 64635 64695 64755 64815 64875 64935 64995 65055 65115 65175 65235 65295 65355 65415 65475 65535 65595 65655 65715 65775 65835 65895 65955 66015 66075 66135 66195 66255 66315 66375 66435 66495 66555 66615 66675 66735 66795 66855 66915 66975 67035 67095 67155 67215 67275 67335 67395 67455 67515 67575 67635 67695 67755 67815 67875 67935 67995 68055 68115 68175 68235 68295 68355 68415 68475 68535 68595 68655 68715 68775 68835 68895 68955 69015 69075 69135 69195 69255 69315 69375 69435 69495 69555 69615 69675 69735 69795 69855 69915 69975 70035 70095 70155 70215 70275 70335 70395 70455 70515 70575 70635 70695 70755 70815 70875 70935 70995 71055 71115 71175 71235 71295 71355 71415 71475 71535 71595 71655 71715 71775 71835 71895 71955 72015 72075 72135 72195 72255 72315 72375 72435 72495 72555 72615 72675 72735 72795 72855 72915 72975 73035 73095 73155 73215 73275 73335 73395 73455 73515 73575 73635 73695 73755 73815 73875 73935 73995 74055 74115 74175 74235 74295 74355 74415 74475 74535 74595 74655 74715 74775 74835 74895 74955 75015 75075 75135 75195 75255 75315 75375 75435 75495 75555 75615 75675 75735 75795 75855 75915 75975 76035 76095 76155 76215 76275 76335 76395 76455 76515 76575 76635 76695 76755 76815 76875 76935 76995 77055 77115 77175 77235 77295 77355 77415 77475 77535 77595 77655 77715 77775 77835 77895 77955 78015 78075 78135 78195 78255 78315 78375 78435 78495 78555 78615 78675 78735 78795 78855 78915 78975 79035 79095 79155 79215 79275 79335 79395 79455 79515 79575 79635 79695 79755 79815 79875 79935 79995 80055 80115 80175 80235 80295 80355 80415 80475 80535 80595 80655 80715 80775 80835 80895 80955 81015 81075 81135 81195 81255 81315 81375 81435 81495 81555 81615 81675 81735 81795 81855 81915 81975 82035 82095 82155 82215 82275 82335 82395 82455 82515 82575 82635 82695 82755 82815 82875 82935 82995 83055 83115 83175 83235 83295 83355 83415 83475 83535 83595 83655 83715 83775 83835 83895 83955 84015 84075 84135 84195 84255 84315 84375 84435 84495 84555 84615 84675 84735 84795 84855 84915 84975 85035 85095 85155 85215 85275 85335 85395 85455 85515 85575 85635 85695 85755 85815 85875 85935 85995 86055 86115 86175 86235 86295 86355 86415 86475 86535 86595 86655 86715 86775 86835 86895 86955 87015 87075 87135 87195 87255 87315 87375 87435 87495 87555 87615 87675 87735 87795 87855 87915 87975 88035 88095 88155 88215 88275 88335 88395 88455 88515 88575 88635 88695 88755 88815 88875 88935 88995 89055 89115 89175 89235 89295 89355 89415 89475 89535 89595 89655 89715 89775 89835 89895 89955 90015 90075 90135 90195 90255 90315 90375 90435 90495 90555 90615 90675 90735 90795 90855 90915 90975 91035 91095 91155 91215 91275 91335 91395 91455 91515 91575 91635 91695 91755 91815 91875 91935 91995 92055 92115 92175 92235 92295 92355 92415 92475 92535 92595 92655 92715 92775 92835 92895 92955 93015 93075 93135 93195 93255 93315 93375 93435 93495 93555 93615 93675 93735 93795 93855 93915 93975 94035 94095 94155 94215 94275 94335 94395 94455 94515 94575 94635 94695 94755 94815 94875 94935 94995 95055 95115 95175 95235 95295 95355 95415 95475 95535 95595 95655 95715 95775 95835 95895 95955 96015 96075 96135 96195 96255 96315 96375 96435 96495 96555 96615 96675 96735 96795 96855 96915 96975 97035 97095 97155 97215 97275 97335 97395 97455 97515 97575 97635 97695 97755 97815 97875 97935 97995 98055 98115 98175 98235 98295 98355 98415 98475 98535 98595 98655 98715 98775 98835 98895 98955 99015 99075 99135 99195 99255 99315 99375 99435 99495 99555 99615 99675 99735 99795 99855 99915 99975 100035 100095 100155 100215 100275 100335 100395 100455 100515 100575 100635 100695 100755 100815 100875 100935 100995 101055 101115 101175 101235 101295 101355 101415 101475 101535 101595 101655 101715 101775 101835 101895 101955 102015 102075 102135 102195 102255 102315 102375 102435 102495 102555 102615 102675 102735 102795 102855 102915 102975 103035 103095 103155 103215 103275 103335 103395 103455 103515 103575 103635 103695 103755 103815 103875 103935 103995 104055 104115 104175 104235 104295 104355 104415 104475 104535 104595 104655 104715 104775 104835 104895 104955 105015 105075 105135 105195 105255 105315 105375 105435 105495 105555 105615 105675 105735 105795 105855 105915 105975 106035 106095 106155 106215 106275 106335 106395 106455 106515 106575 106635 106695 106755 106815 106875 106935 106995 107055 107115 107175 107235 107295 107355 107415 107475 107535 107595 107655 107715 107775 107835 107895 107955 108015 108075 108135 108195 108255 108315 108375 108435 108495 108555 108615 108675 108735 108795 108855 108915 108975 109035 109095 109155 109215 109275 109335 109395 109455 109515 109575 109635 109695 109755 109815 109875 109935 109995 110055 110115 110175 110235 110295 110355 110415 110475 110535 110595 110655 110715 110775 110835 110895 110955 111015 111075 111135 111195 111255 111315 111375 111435 111495 111555 111615 111675 111735 111795 111855 111915 111975 112035 112095 112155 112215 112275 112335 112395 112455 112515 112575 112635 112695 112755 112815 112875 112935 112995 113055 113115 113175 113235 113295 113355 113415 113475 113535 113595 113655 113715 113775 113835 113895 113955 114015 114075 114135 114195 114255 114315 114375 114435 114495 114555 114615 114675 114735 114795 114855 114915 114975 115035 115095 115155 115215 115275 115335 115395 115455 115515 115575 115635 115695 115755 115815 115875 115935 115995 116055 116115 116175 116235 116295 116355 116415 116475 116535 116595 116655 116715 116775 116835 116895 116955 117015 117075 117135 117195 117255 117315 117375 117435 117495 117555 117615 117675 117735 117795 117855 117915 117975 118035 118095 118155 118215 118275 118335 118395 118455 118515 118575 118635 118695 118755 118815 118875 118935 118995 119055 119115 119175 119235 119295 119355 119415 119475 119535 119595 119655 119715 119775 119835 119895 119955 120015 120075 120135 120195 120255 120315 120375 120435 120495 120555 120615 120675 120735 120795 120855 120915 120975 121035 121095 121155 121215 121275 121335 121395 121455 121515 121575 121635 121695 121755 121815 121875 121935 121995 122055 122115 122175 122235 122295 122355 122415 122475 122535 122595 122655 122715 122775 122835 122895 122955 1
```











Wed Mar 17 09:43:28 1999

```

PT creating animal models for drug screening
PS Example 1: Fig 9b-g: 82pp; English.
CC Vector pTet-Splice (T49876) is utilized in the construction of
CC autoregulatory vector plasmid pTet-trak (T49877). Splice-PA was
CC made by ligating the SV40 small T antigen intervening sequence and
CC the SV40 early polyA sequence into pBSK11+. The XhoI-SalI fragment
CC of pUHC13-3 (contg. 7 copies of the tet operator upstream of minimal
CC promoter Tetp) was cloned upstream of the splice/polyA sequence of
CC pSplice-PA to form pTet-Splice. A modified tetracycline
CC transactivator (tTAK) gene was cloned into pTet-Splice to form pTet-
CC trak. The construct provides rapid, reversible control of gene
CC expression in eukaryotic cells or transgenic animals, e.g. for
CC heterologous protein, RNA or antisense sequence prodn., or for
CC creating animal models of human disease.
CC Sequence 5178 BP: 1383 A; 1169 C; 1240 G; 1386 T;
SQ
Query Match 66.0%, Score 22.2, EB 27, Length 5178;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 2957 ctcagagggggcccggtaccacgcttttgccttttagtgaggggttaattgcgcgtt 3016
QY 6 CTCAGAGGGGGCCCGGTACCACGCTTTTGCCTTTAGTGAGGGTAAATTGCGCGCTT 65
Db 3017 ggcgttaattcgttcagctgtttcctctgtgtgaaattttatccgctcacattccaca 3076
QY 66 GCGTAAATCATGGTCATAGCTGTTTCCTGTGTGAAATTTATCCGCTCACAAATCCACA 125
Db 3077 caactacagccgagcagcataaagtaagcctggggtgcctaatgaatgaatgaact 3136
QY 126 CAACATACAGCCGGAACATAAAGTGTAAAGCTTGGGTGCTTAATGAGTGAAT 185
Db 3137 cacattattgcttgcctcactgcccgcctttccacgtcgggaaacactgtctctgcagct 3196
QY 186 CACATTAAATGGCTTGGCTGCTCACTGCTGCTTTCACGTGCGGAAACCTGTCTGACGCT 245
Db 3197 gaattataaatacggcgaacacgaggggagagcggtttcgtattgagcgctctctccg 3256
QY 246 GCATTAAATGAATCGGCCAACGCGCGGGGAGAGCGGTTTCTGTAATGGGCTTTCGCG 305
Db 3257 ttctcgcctcactgactcgcctgcctgcctgcctgcctgcctgcctgcctgcctgcct 3316
QY 306 TTCTCGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 365
Db 3317 ctcaaggggtaatacgtttatccacagaaatcagggtataacgcaggaagaaacatgtg 3376
QY 366 CTCAAAGGCGGTAAATACGGTTATCCACAGATCAGGGGATAACGCAGGAAAGAAATGTG 425
Db 3377 agcaaaaggccagcaaaaggccaggaacccgtataaaagccgcctgttgctgagcttttcca 3436
QY 426 AGCAAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAGAGCCGCTTTCCTGGCTTTTTCCTCA 485
Db 3437 taggtccgccccctgacgagcatcacaaaaatcacactcaagtcagagtgagcgaaga 3496
QY 486 TAGGTCCGCCCCCTGACGAGCATACAAAAAAGCAAGCTCAATCAGAGGTGGGAAA 545
Db 3497 ccgacagagcactaaagataccagagcgttttccccttggaaactccctgagctgtctcc 3556
QY 546 CCGGACAGGACTATAAAGATACAGAGGCTTTCGCGCTGGAAGCTCCCTGCTGCGCTCC 605
Db 3557 tcttcgacctgacgcttaccgagatcacctgtcgcctttctcccttcgggagagctggc 3616
QY 606 TGTTCCGACCCCTGCGCTTTACCGGATACCTGTCCGCTTTCCTGCTGGGAGGTGGC 665
Db 3617 gctttctcagctcagcgtgttaggtatctcagttcagttcagttcagttcagttcagtt 3676
QY 666 GCTTCTCATAGCTCACGCTTAGGTATCTCAGTTCGCTGAGGTGAGGTGCTTCCGCTCCAGCT 725
Db 3677 ggcctgtgtgacgaaccccccttcagccgacgcgtgcgccttaccgttaccgttaccgt 3736
QY 726 GGCCTGTGTGACGAACCCCCCTTCAGCCGCACTGCTTCCGCTTATCCGCTTATCCGCT 785
Db 3737 tcttgagtcacaccccggttaagacacgacttatcgccacttctctctctctctctctct 3796

```

RESULT 15

```

ID T49875 standard; DNA; 5178 BP.
AC T49876;
PT 28-MAR-1997 (first entry)
DE pTet-Splice.
KW pTet-Splice; pTet-tTAK; tetracycline transactivator; promoter;
KW gene expression; transgenic animal; animal model; drug screening;
KW vector; ds
OS Synthetic.
FH Key Location/Qualifiers
FT misc_rna complement (2591)
FT FT /*tag= a
FT FT /note= "putative start site of trxn"
FT FT complement (2616 2622)
FT FT /*tag= b
FT tata_signal
FT WO9640946-A1
PD 19-DEC-1996
PR 07-JUN-1996: 010109
PR 07-JUN-1995: US-474169.
PA (UYA ) UNIV YALE.
PI Schatz DG;
CR WP1 97-07273/07
PI Nucleic acid encoding tetracycline transactivator fusion protein.
PI provides rapid and reversible control of gene expression, e.g. for
PT

```







Release 3.1A John F. Collins, Biocomputing Research Unit  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	664	19.7	711	26	AGC04608	Homo sapiens genomic	D 0.00e+00
2	664	19.7	711	27	AGC10947	Homo sapiens genomic	D 0.00e+00
3	657	19.5	721	27	AGC11001	Homo sapiens genomic	D 0.00e+00
4	657	19.5	721	26	AGC04652	Homo sapiens genomic	D 0.00e+00
5	635	18.9	718	26	AGC04363	Homo sapiens genomic	D 0.00e+00
6	635	18.9	718	27	AGC10489	Homo sapiens genomic	D 0.00e+00
7	630	18.7	698	26	AGC03787	Homo sapiens genomic	D 0.00e+00
8	630	18.7	698	27	AGC09976	Homo sapiens genomic	D 0.00e+00
9	624	18.5	696	27	AGC03765	Homo sapiens genomic	D 0.00e+00
10	624	18.5	696	26	AGC04576	Homo sapiens genomic	D 0.00e+00
11	524	18.5	720	28	AGC13858	Homo sapiens genomic	D 0.00e+00
12	624	18.5	720	26	AGC04062	Homo sapiens genomic	D 0.00e+00
13	619	18.4	747	26	AGC07052	Homo sapiens genomic	D 0.00e+00

## ALIGNMENTS

BASE COUNT	$169 \pm$	$188 \pm$	$172 \pm$	$192 \pm$
	$2.14$	$2.14$	$2.14$	$2.14$

```

Query Match      19.78;  Score 664;  DB 26;  Length 711;
Best Local Similarity 99.6%;  Pred. NO. 0.00e+00;
Matches 682;  Conservative 0;  Mismatches 0;  Indels

```

Best Local Similarity	99.6%;	Pred. No. 0.00e+00;	Indels	0.	Gaps	0.
			Mismatches	0.		



```

REFERENCE
1 (bases 1 to 721)
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
AUTHORS
Homo sapiens genomic DNA, chromosome 21q
TITLE
Published Only in Database (1998) In press
JOURNAL
2 (bases 1 to 721)
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
AUTHORS
Direct Submission
TITLE
Submitted (29-JUL-1998) to the DDBJ/EMBL/GenBank databases.
JOURNAL
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@ngc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
FEATURES
Location/Qualifiers
1..721
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="879G12X91"
/map="21q"
BASE COUNT
195 a 163 c 175 g 179 t 9 others
ORIGIN
Query Match 19.5%: Score 657, DB 27, Length 721.
Best Local Similarity 98.7%: Pred. No. 0.00e+00;
Matches 683; Conservative 0; Mismatches 5; Indels 4; Gaps 4;
Db 32 CATCTTACGGATGGCATGACAGTAAGAGAATTTATGCAGTGTGTCATTAACCATGAGTGAT 91
Cp 1772 CATCTTACGGATGGCATGACAGTAAGAGAATTTATGCAGTGTGTCATTAACCATGAGTGAT 1713
Db 92 AACACTGCGCGCAACTTACTTCTGACACAGTATCGGAGGACCGAAGGAGCTAACCGCTTTT 151
Cp 1712 AACACTGCGCGCAACTTACTTCTGACACAGTATCGGAGGACCGAAGGAGCTAACCGCTTTT 1653
Db 152 TTGCACAAATCGGGGATCATGTAACTCGCTTGTATGCTTGGGAAACCGAGCTGAATGAA 211
Cp 1652 TTGCACAAATCGGGGATCATGTAACTCGCTTGTATGCTTGGGAAACCGAGCTGAATGAA 1593
Db 212 GCGATACCAACAGAGAGCGGTGACACACGATGCTGTAGCAATGGCAACAGCTTGGCG 271
Cp 1592 GCGATACCAACAGAGAGCGGTGACACACGATGCTGTAGCAATGGCAACAGCTTGGCG 1533
Db 272 AAACCTAATTAACCTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 331
Cp 1532 AAACCTAATTAACCTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 1473
Db 332 GAGCGGATTAACCTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 391
Cp 1472 GAGCGGATTAACCTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 1413
Db 392 GCGATACCAACAGAGAGCGGTGACACACGATGCTGTAGCAATGGCAACAGCTTGGCG 451
Cp 1412 GCGATACCAACAGAGAGCGGTGACACACGATGCTGTAGCAATGGCAACAGCTTGGCG 1353
Db 452 SATCGTAAGCTTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 511
Cp 1352 SATCGTAAGCTTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 1293
Db 512 GAGCGGATTAACCTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 570
Cp 1292 GAGCGGATTAACCTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 1233
Db 571 GAGCGGATTAACCTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 629
Cp 1232 GAGCGGATTAACCTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 1173
Db 629 ATCTAGTGAAGATCTTTTTCATATCTGATGATGATGATGATGATGATGATGATGATGATG 687
Cp 1172 ATCTAGTGAAGATCTTTTTCATATCTGATGATGATGATGATGATGATGATGATGATG 1113
Db 687 TTCCACTGAGCGTCAGACCCCGGTAGAAAAAT 719
Cp 1112 TTCCACTGAGCGTCAGACCCCGGTAGAAAAAT 1081

```

```

4
RESULT
LOCUS
DEFINITION
Homo sapiens genomic DNA, 21q region, clone: 879G12X91, genomic
survey sequence.
ACCESSION
AG004562
NID
g2826191
KEYWORDS
GSS.
SOURCE
Homo sapiens DNA, clone: 879G12X91.
ORGANISM
Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 721)
Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
AUTHORS
Homo sapiens genomic DNA, chromosome 21q
TITLE
Published Only in Database (1998) In press
JOURNAL
2 (bases 1 to 721)
Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
AUTHORS
Direct Submission
TITLE
Submitted (30-JAN-1998) to the DDBJ/EMBL/GenBank databases.
JOURNAL
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@ngc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
FEATURES
Location/Qualifiers
1..721
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="879G12X91"
/map="21q"
BASE COUNT
195 a 163 c 175 g 179 t 9 others
ORIGIN
Query Match 19.5%: Score 657, DB 26, Length 721.
Best Local Similarity 98.7%: Pred. No. 0.00e+00;
Matches 683; Conservative 0; Mismatches 5; Indels 4; Gaps 4;
Db 32 CATCTTACGGATGGCATGACAGTAAGAGAATTTATGCAGTGTGTCATTAACCATGAGTGAT 91
Cp 1772 CATCTTACGGATGGCATGACAGTAAGAGAATTTATGCAGTGTGTCATTAACCATGAGTGAT 1713
Db 92 AACACTGCGCGCAACTTACTTCTGACACAGTATCGGAGGACCGAAGGAGCTAACCGCTTTT 151
Cp 1712 AACACTGCGCGCAACTTACTTCTGACACAGTATCGGAGGACCGAAGGAGCTAACCGCTTTT 1653
Db 152 TTGCACAAATCGGGGATCATGTAACTCGCTTGTATGCTTGGGAAACCGAGCTGAATGAA 211
Cp 1652 TTGCACAAATCGGGGATCATGTAACTCGCTTGTATGCTTGGGAAACCGAGCTGAATGAA 1593
Db 212 GCGATACCAACAGAGAGCGGTGACACACGATGCTGTAGCAATGGCAACAGCTTGGCG 271
Cp 1592 GCGATACCAACAGAGAGCGGTGACACACGATGCTGTAGCAATGGCAACAGCTTGGCG 1533
Db 272 AAACCTAATTAACCTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 331
Cp 1532 AAACCTAATTAACCTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 1473
Db 332 GAGCGGATTAACCTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 391
Cp 1472 GAGCGGATTAACCTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 1413
Db 392 GCGATACCAACAGAGAGCGGTGACACACGATGCTGTAGCAATGGCAACAGCTTGGCG 451
Cp 1412 GCGATACCAACAGAGAGCGGTGACACACGATGCTGTAGCAATGGCAACAGCTTGGCG 1353
Db 452 SATCGTAAGCTTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 511
Cp 1352 SATCGTAAGCTTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 1293
Db 512 GAGCGGATTAACCTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 570
Cp 1292 GAGCGGATTAACCTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 1233
Db 571 GAGCGGATTAACCTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 629
Cp 1232 GAGCGGATTAACCTGGCAACTACTTACTTCTGCGGCAACAAATTAATAGACTGGATG 1173
Db 629 ATCTAGTGAAGATCTTTTTCATATCTGATGATGATGATGATGATGATGATGATGATGATG 687
Cp 1172 ATCTAGTGAAGATCTTTTTCATATCTGATGATGATGATGATGATGATGATGATGATG 1113
Db 687 TTCCACTGAGCGTCAGACCCCGGTAGAAAAAT 719
Cp 1112 TTCCACTGAGCGTCAGACCCCGGTAGAAAAAT 1081

```



Db	141	CATCAAAAATCGACGCTCAAGTCAGAGTGGGAAACCGACAGACTATAAGATAC	200
Qy	508	CATCAAAAATCGACGCTCAAGTCAGAGTGGGAAACCGACAGACTATAAGATAC	567
Db	201	CAGCGGTTTCGGGCTTGGAAAGCTCCCTCGTGGGCTTCTGTTCGAGCCCTGGCGCTTACC	260
Qy	508	CAGCGGTTTCGGGCTTGGAAAGCTCCCTCGTGGGCTTCTGTTCGAGCCCTGGCGCTTACC	527
Db	261	CGATAGCTATCGGCTTTCTCCCTTCGGGAAAGCTGGCGCTTTCTCATAGCTCAGCGTGT	320
Qy	629	CGATAGCTATCGGCTTTCTCCCTTCGGGAAAGCTGGCGCTTTCTCATAGCTCAGCGTGT	687
Ld	321	AGGATCTCAAGTTCGGTGTAGGTGCTTCCTCCAGCTGGGCTGTGTGCAGAACCCGCC	380
Qy	588	AGGATCTCAAGTTCGGTGTAGGTGCTTCCTCCAGCTGGGCTGTGTGCAGAACCCGCC	747
Db	381	GTACGCGGCAAGCGCTTCGGGCTTATCGGTAAGTATCGCTTTATCGGTAAGTATCGCT	440
Qy	748	GTACGCGGCAAGCGCTTCGGGCTTATCGGTAAGTATCGCTTTATCGGTAAGTATCGCT	807
Db	441	CACGACATTATCGGCACTGGCAGAGCCACTGGTAACAGGATTAGTAGACNAGGTATGTA	500
Qy	808	CACGACATTATCGGCACTGGCAGAGCCACTGGTAACAGGATTAGTAGACNAGGTATGTA	867
Db	501	GGCGGCGTATACAGAGTCTCTGAAGTGGTGGCTTAAGTCGGCTAACACTAGAAGAACAAG	560
Qy	868	GGCGGCGTATACAGAGTCTCTGAAGTGGTGGCTTAAGTCGGCTAACACTAGAAGAACAAG	925
Db	561	TATTTGGTATCTGGGCTCTGCTGAAGCCAGTTANCTTTCGGAAGAAAGTGTGTAGCTCT	620
Qy	924	TATTTGGTATCTGGGCTCTGCTGAAGCCAGTTANCTTTCGGAAGAAAGTGTGTAGCTCT	984
Db	621	TGATGGGCGAACAACACCGCTGGTAGCG-TGTTTTTTTGTTCGAGGACGACAGATA	679
Qy	995	TGATGGGCGAACAACACCGCTGGTAGCG-TGTTTTTTTGTTCGAGGACGACAGATT	1044
Db	680	ATCGCGCAGAAAAAAGGATCTCAAGA-GATC-TTTGATCTT	718
Qy	1045	ATCGCGCAGAAAAAAGGATCTCAAGAAGATCTTTGATCTT	1085

RESULT	7	AC003787	698 bp	DNA	GSS	07-JAN-1998
DEFINITION		Homo sapiens genomic DNA, 21q region, clone: T485XN, genomic survey				
LOCUS		sequence.				
ACCESSION		AC003787				
NID		q2754689				
KEYWORDS		GSS.				
SOURCE		Homo sapiens DNA, clone:T485XN.				
ORGANISM		Homo sapiens				
		Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;				
		Primates; Catarrhini; Hominidae; Homo.				
REFERENCE		1 (bases 1 to 698)				
AUTHORS		Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.				
TITLE		Homo sapiens genomic DNA, chromosome 21q				
JOURNAL		Published Only in Database (1998) in press				
REFERENCE		2 (bases 1 to 698)				
AUTHORS		Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.				
TITLE		Direct Submission				
JOURNAL		Submitted (07-JAN-1998) to the DDBJ/EMBL/Genbank databases.				
		Masahira Hattori, Kitasato University, Department of Science, JST				
		Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan				
		(E-mail:hattori@nc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,				
		Fax:0427-78-9561)				
FEATURES		Location/Qualifiers				
		1	698			
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
		/chromosome="21"				
		/clone="T485XN"				
		/map="21q"				
		189 a	153 c	171 g	175 t	10 others
		BASE COUNT				
		ORIGIN				

Query Match	18.7%	Score 630.	DB 26:	Length 598:
Best Local Similarity	99.0%:	Pred. No. 0.00e+00:		
Matches	661:	Mismatches	2:	Gaps 5:
		Conservative		
Db	31	CTTAGCGATGGCATGACAGTANGAGTAATTCAGTGCTGCCATACCATCAGTGTATAAC	90	
Cp	1769	CTTAGCGATGGCATGACAGTAAGAATAATTATGCAGTCTGCCATAAACATCAGTGTAAAC	1710	
Db	91	ACTGGCGCCCACTTACTTTGACAACGATCGGAGACCGAGAGAGCTAACCGCTTTITG	150	
Cp	1709	ACTGGCGCCCACTTACTTTGACAACGATCGGAGACCGAGAGAGCTAACCGCTTTITG	1650	
Db	151	CACACATCGGGGAATCATGTAACTCGGCTTGATTCGTTGGAAACGAGCTGAATGAAGCC	210	
Cp	1649	CACAAUATGGGGATCATGTANTCTGGCTTGATTCGTTGGAAACGAGCTGAATGAAGCC	1590	
Db	211	AATCAAAGSAGSAGSGCTGATACCACAGATGCTTTAGCAATGGCACACACGTTGGCGAAA	270	
Cp	1589	ATACCAAGCAGCGGTGTCACACCAAGATGCTGTAGCAATGGCAACACGTTGGCGAAA	1530	
Db	271	CTATTAACTGGCCAAGTACTTACTATCTAGCTTTCCGCGCAACAATTAATAGACTGGATGGAS	330	
Cp	1529	CTATTAACTGGCCAAGTACTTACTATCTAGCTTTCCGCGCAACAATTAATAGACTGGATGGAG	1470	
Db	331	CGCGATAAATCTTCAGAGACACATTCGCGCTCGGCCCCTTCGCGTGGCTGGTTATGCT	390	
Cp	1458	GGGAIANAAGTTCAGAGACACATTCGCGCTCGGCCCCTTCGCGTGGCTGGTTATGCT	1410	
Db	391	GATAAATCTGGAGCCGCTGAGTGGCTCGCGGTATCATTCGAGCACCTGGGGCCAGAT	450	
Cp	1409	GATAAATCTGGAGCCGCTGAGTGGCTCGCGGTATCATTCGAGCACCTGGGGCCAGAT	1350	
Db	451	GGTAAGCCCTCCCGTATCGTAGTTATCTACAGACGGG-AGTCAGGCACTATGATGA	509	
Cp	1349	GGTAAGCCCTCCCGTATCGTAGTTATCTACAGACGGGAGTCAGGCACTATGATGA	1290	
Db	510	CGAAATAGACAGATTCGCTGAGTAGGTGCCTCAGTCATTAAAGCAATGGTAACCTGACGA	569	
Cp	1289	CGAAATAGACAGAT-CGCTGAGATAGGTGCCTCAGTCATTAAAGCAATGGTAACCTGACGA	1231	
Db	570	CCAAGTTTACTC-TATATACCTTTAGATTGAATTTAAACCTTCATTTTTTAATAAGGAT	628	
Cp	1230	CCAAGTTTACTCATATATACCTTTAGATTGAATTTAAACCTTCATTTTTTAATAAGGAT	1171	
Db	629	CTAGGTGAAGATCNNTTT-GATAATCTCATGACCAAAATCCCTTAAAGCTGAGTTTCGTT	687	
Cp	1170	CTAGGTGAAGATCCCTTTTGATAATCATGACCAAAATCCCTTAAAGCTGAGTTTCGTT	1111	
Db	688	TCCACTGA	695	
Cp	1110	-CCACTGA	1104	

RESULT	8	AG009976	698 bp	DNA	GSS	06-JUL-1998
LOCUS						
DEFINITION				Homo sapiens genomic DNA, 21q region, clone: T485XN, genomic survey sequence.		
ACCESSION				AG009976		
NID				g3289962		
KEYWORDS				GSS.		
SOURCE				Homo sapiens DNA, clone:T485XN.		
ORGANISM				Homo sapiens		
				Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;		
				Primates; Catarrhini; Hominoidea; Homo.		
REFERENCE				1 (bases 1 to 698)		
AUTHORS				Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.		
JOURNAL				Homo sapiens genomic DNA, chromosome 21q		
TITLE				Published only in Database (1998) In press		
REFERENCE				2 (bases 1 to 698)		
AUTHORS				Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.		
TITLE				Direct submission		
JOURNAL				Submitted (06-JUL-1998) to the DDBJ/EMBL/Genbank databases.		

Masahira Hattori, Kitasato University, Department of Science, JST  
Sequencing Laboratory, 1-15-1, Sakurabara 228, Japan  
(E-mail:hattori@ims.u-tokyo.ac.jp, Tel:0427-78-5932,  
Fax:0427-78-9561)

## FEATURES

Location/Qualifiers

SOURCE

1..696

Organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="21"

/accession="U480975"

/map="21q"

BASE COUNT 189 a 153 c 171 g 175 t 10 others

## ORIGIN

Query Match 18.7% Score 640; DB 27; Length 696;

Best Local Similarity 99.1% Pred No. 0.00e+00;

Matches 661; Conservative 0; Mismatches 2; Indels 5; Gaps 5;

Db 31 GTTACGATGGGATACATTAAGCAAAATTTATGATGTTGGTATAACCATGAGTGATAAC 90  
|||||  
Cp 1769 GTTACGATGGGATACATTAAGCAAAATTTATGATGTTGGTATAACCATGAGTGATAAC 1710  
Db 91 ACTGGGTAATTAATTTGATTAAGCAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 150  
|||||  
Cp 1709 ACTGGGTAATTAATTTGATTAAGCAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1650  
Db 151 CAAACATGAGGATGATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 210  
|||||  
Cp 1649 CAAACATGAGGATGATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1590  
Db 211 ATACCAAAATGACCAAGGATGATTAATTTGATGATGATGATGATGATGATGATGATGATG 270  
|||||  
Cp 1589 ATACCAAAATGACCAAGGATGATTAATTTGATGATGATGATGATGATGATGATGATGATG 1530  
Db 271 TATTACTGGGAATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 330  
|||||  
Cp 1524 TATTACTGGGAATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1470  
Db 331 GGTGATAAATTTGATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 390  
|||||  
Cp 1469 GGTGATAAATTTGATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1410  
Db 441 GATAAATCTGGAGGATGATTAATTTGATGATGATGATGATGATGATGATGATGATGATG 450  
|||||  
Cp 1409 GATAAATCTGGAGGATGATTAATTTGATGATGATGATGATGATGATGATGATGATGATG 1350  
Db 451 GGTAAAGGCTGGGATGATTAATTTGATGATGATGATGATGATGATGATGATGATGATG 509  
|||||  
Cp 1349 GGTAAAGGCTGGGATGATTAATTTGATGATGATGATGATGATGATGATGATGATGATG 1290  
Db 520 GAAATAGATTAATTTGATTAATTTGATGATGATGATGATGATGATGATGATGATGATG 569  
|||||  
Cp 1289 GAAATAGATTAATTTGATTAATTTGATGATGATGATGATGATGATGATGATGATGATG 1231  
Db 570 GAAAGTTTACTGATTAATTTGATTAATTTGATGATGATGATGATGATGATGATGATGATG 628  
|||||  
Cp 1230 GAAAGTTTACTGATTAATTTGATTAATTTGATGATGATGATGATGATGATGATGATGATG 1171  
Db 629 GTATGTTGAATGATGATTTGATTAATTTGATGATGATGATGATGATGATGATGATGATG 687  
|||||  
Cp 1172 GTATGTTGAATGATGATTTGATTAATTTGATGATGATGATGATGATGATGATGATGATG 1111  
Db 688 TTAATGTA 695  
|||||  
Cp 1110 TTAATGTA 1104

RESIDUE AC009765 646 bp DNA 05-JUL-1998  
FEATURES  
SOURCE Homo sapiens genomic DNA; 21q region; clone: P84SfN15; genomic  
survey sequence.  
AC009765  
NOTE G328975)

## KEYWORDS

Homo sapiens DNA, clone:P84SfN15.

## SOURCE

Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Hominoidea; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 696)  
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakabe,Y.

## AUTHORS

Homo sapiens genomic DNA, chromosome 21q

## JOURNAL

Published only in Database (1998) In press

## REFERENCE

2 (bases 1 to 696)  
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakabe,Y.

## AUTHORS

Direct Submission

## TITLE

Submitted (06-09-1998) to the EMBL/GenBank/DDBJ databases.

## JOURNAL

Masahira Hattori, Kitasato University, Department of Science, JST

Sequencing Laboratory, Kitasato 1-15-1, Sakurabara 228, Japan  
(E-mail:hattori@ims.u-tokyo.ac.jp, Tel:0427-78-5932,  
Fax:0427-78-9561)

## FEATURES

Location/Qualifiers

SOURCE

1..696

Organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="21"

/accession="P84SfN15"

/map="21q"

BASE COUNT 183 a 165 c 185 g 162 t 1 others

## ORIGIN

Query Match 18.5% Score 624; DB 27; Length 696;

Best Local Similarity 99.1% Pred No. 0.00e+00;

Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 31 GTTTTGGTCAGCAAGAAAGCTGGTAAATTAATAATGATTAATAATGATTAATAATGATTA 90  
|||||  
Cp 2039 GTTTTGGTCAGCAAGAAAGCTGGTAAATTAATAATGATTAATAATGATTAATAATGATTA 1989  
Db 91 CGAGTGGGTTACATGAACTGGTAAATTAATAATGATTAATAATGATTAATAATGATTAATA 150  
|||||  
Cp 1979 CGAGTGGGTTACATGAACTGGTAAATTAATAATGATTAATAATGATTAATAATGATTAATA 1920  
Db 151 GAAGAACGTTTGGTAAATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTAATA 210  
|||||  
Cp 1919 GAAGAACGTTTGGTAAATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTAATA 1869  
Db 211 GTATTGAGGCTGGGATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTAATA 270  
|||||  
Cp 1859 GTATTGAGGCTGGGATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTAATA 1809  
Db 271 GTTGAATGATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTA 330  
|||||  
Cp 1799 GTTGAATGATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTA 1749  
Db 331 TGCAGTGGCTGGCATAACCATGAGTGATTAATAATGATTAATAATGATTAATAATGATTAATA 390  
|||||  
Cp 1739 TGCAGTGGCTGGCATAACCATGAGTGATTAATAATGATTAATAATGATTAATAATGATTAATA 1689  
Db 391 GAGACGCTGGGATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTAATAATGAT 450  
|||||  
Cp 1679 GAGACGCTGGGATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTAATAATG 1629  
Db 451 GATCGTTGGTAATGATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTAATAAT 510  
|||||  
Cp 1619 GATCGTTGGTAATGATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTAATA 1569  
Db 511 GGCGTTGAGCAATGGCAATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTAATA 570  
|||||  
Cp 1560 GGCGTTGAGCAATGGCAATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTA 1510  
Db 571 GTTCTGGGATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTA 630  
|||||  
Cp 1501 GTTCTGGGATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTA 1451  
Db 631 GTTCTGGGATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTAATAATGATTA 690  
|||||

Chromosome	Position	Sequence	Source	Accession	Keywords	Organism	Reference	Title	Journal	Author	Year	Page
1	1619	GATCGTTGGACCGGAGCTGAATGAGCCATACAAACGACGACGTGA-CACACGAT	1561									
1	511	GGCTGTAGCAATGGCAACAGGTTGGCAAACTATTAACTGGCAACTACTTACTCTAS	570									
1	1560	G-CCTGTAGCAATGGCAACAGGTTGGCAAACTATTAACTGGCAACTACTTACTCTAG	1502									
1	571	CTTCCGGGCAACAATTAATAGACTGGATGGAGGGGATGAAGTTGCGAGCCACTTCTG	630									
1	1501	CTTCCGGG-CAACAATTAATAGACTGGATGGAGGGGATGAAGTTGCGAGCCACTTCTG	1443									
1	631	CGCTGGCCCTTCGCGCTGGCTGTTTATTGCTGATAAACTGAGAGCGGAGC-TG	689									
1	1442	CGCTGGCCCTTCGCGCTGG-CTGGTCTG-CTGGTATTGCTGATAAACTGAGAGCGGAGC-TG	1385									
1	690	GGTCTC 695										
1	1384	GGTCTC 1379										
1	1442	CGCTGGCCCTTCGCGCTGG-CTGGTCTG-CTGGTATTGCTGATAAACTGAGAGCGGAGC-TG	1385									
1	690	GGTCTC 695										
1	1384	GGTCTC 1379										
1	1442	CGCTGGCCCTTCGCGCTGG-CTGGTCTG-CTGGTATTGCTGATAAACTGAGAGCGGAGC-TG	1385									
1	690	GGTCTC 695										
1	1384	GGTCTC 1379										
1	1442	CGCTGGCCCTTCGCGCTGG-CTGGTCTG-CTGGTATTGCTGATAAACTGAGAGCGGAGC-TG	1385									
1	690	GGTCTC 695										
1	1384	GGTCTC 1379										
1	1442	CGCTGGCCCTTCGCGCTGG-CTGGTCTG-CTGGTATTGCTGATAAACTGAGAGCGGAGC-TG	1385									
1	690	GGTCTC 695										
1	1384	GGTCTC 1379										
1	1442	CGCTGGCCCTTCGCGCTGG-CTGGTCTG-CTGGTATTGCTGATAAACTGAGAGCGGAGC-TG	1385									
1	690	GGTCTC 695										
1	1384	GGTCTC 1379										
1	1442	CGCTGGCCCTTCGCGCTGG-CTGGTCTG-CTGGTATTGCTGATAAACTGAGAGCGGAGC-TG	1385									
1	690	GGTCTC 695										
1	1384	GGTCTC 1379										
1	1442	CGCTGGCCCTTCGCGCTGG-CTGGTCTG-CTGGTATTGCTGATAAACTGAGAGCGGAGC-TG	1385									
1	690	GGTCTC 695										
1	1384	GGTCTC 1379										
1	1442	CGCTGGCCCTTCGCGCTGG-CTGGTCTG-CTGGTATTGCTGATAAACTGAGAGCGGAGC-TG	1385									
1	690	GGTCTC 695										
1	1384	GGTCTC 1379										
1	1442	CGCTGGCCCTTCGCGCTGG-CTGGTCTG-CTGGTATTGCTGATAAACTGAGAGCGGAGC-TG	1385									
1	690	GGTCTC 695										
1	1384	GGTCTC 1379										
1	1442	CGCTGGCCCTTCGCGCTGG-CTGGTCTG-CTGGTATTGCTGATAAACTGAGAGCGGAGC-TG	1385									
1	690	GGTCTC 695										
1	1384	GGTCTC 1379										
1	1442	CGCTGGCCCTTCGCGCTGG-CTGG										

























wed Mar 17 09:43:33 1999

REFERENCE  
AUTHORS  
TITLE

1 (bases 1 to 9655)  
Mazzarella, R. and Pilia, G.  
Recombination trapping: an 'in vivo' approach to recover cDNAs  
encoded in YACS

JOURNAL  
COMMENT  
FEATURES

Unpublished (1997)  
GSDB:S:127444  
Location/Qualifiers

source  
1..9655

/organism="Cloning vector pGR8"  
/plasmid="Cloning vector pGR8"  
/db\_xref="taxon:64712"

misc\_feature  
1..135

/standard\_name="telomere"

prim\_transcript  
145..2097

/gene="HIS5"

gene  
145..2097

/gene="HIS5"

CDS  
761..1915

/gene="HIS5"

/EC\_number="2.6.1.9"

/codon\_start=1

/transl\_table=11

/product="histidinol phosphate aminotransferase"

/db\_xref="PID:q2358009"

/translation="MWEDLKRIVFRKIYNLEPYRCARDDFTGILLIDANENAHGTPV  
ELSKTNHRYDPHQLFKAAMKYRKNKSSVADPEVAPLTDNLCLGVGSDSIDA  
ITRACVREKILVLPPTYSMSVCANINIEVQCELVDSGFUMDTGDTAVLTIKN  
DSLILMEVTSPNPTGAKTKSLIEKVLONWNLGVVDEAVDSCGSGTAPLVKY  
PNLVLTQTSKSPGLAGITMTYATLARIINAMKAPINISLASEVALKAVQCSN  
LKKMEATSKIIIEEKMLLKELTALDYVCCYVGGDANFLIFIMGNGENAKLIY  
QIATSGVVVFRGNELGSCGLRITVTGTHEENTHLIKYKETLYKLANE"

prim\_transcript  
2138..6708

/gene="LIS2"

gene  
2138..6708

/gene="LIS2"

CDS  
2374..6552

/gene="LIS2"

/EC\_number="1.2.1.31"

/codon\_start=1

/transl\_table=11

/product="alpha-aminoadipate reductase"

/db\_xref="PID:q2358010"

/translation="MINERWIEKLDNPTLSVLPHDFLRPOEPVTKOATVSLQPLQ  
DVPHDSFNKVAVALSWAALYRVIGLDDIVLYIANKILRFN"OPTWSPNELYST  
NNELKLNSTEANFSEDELAEXIQSCODLERTPQLFLAFLENOQKLEDFKHLVDF  
ALNLDTSNAHVLNLIYNSLSNERTIVADQFTQYLTALSRSGNCTIKSLIAS  
SKSLDPTKMLGWCDEVGTHIDIFODNAEAFERTCVETPILNSDKRSFTRDIN  
RTSNVAHLYIKTGIRKGDVMIYSSRGVDMVCMVGLKAGATSVIDPAYPARQI  
LYGVAPRGLIVIRAGOLDLVEDYINDELEIVSRINSIAIQENGIEGKLNQGE  
DVLAPYHDKTRIGVGVGSDSNPTLSFTSGSEGIPKGVGLGRHFLAYEFNMKRFN  
LTDNDKFTMLSGIAHDPIDQDMFTPLFLGALYVPIQDDIGTPGRLEWMSKYGCTV  
RAVSFVEYKSKNDPDLNFKKLDKDMVPMAGKMLNVLVLRNDRTQIGIGIGIYV  
HUTPAGQLTAQTATPPKLRHAFVGLILTKBCLRLQTLAENCRIVNMYGTTETQ  
RAGGLAEGYRGLPELNKEFNMMVEKDHNTLNDKNDSEPRQWFGPRDLRYTGD  
LGRVLPNGDCGCCADKQVIRGFRTELEIDITHISQHPLENTLYRKNADNEPT  
LITVMPREFDPDDLKQFSDVPKEVETPVRKGLLYHLLSKDKIFPLKRLASAM  
PSLIYVMDKPLNPNKGVDPKQLOFTPKMLNVAENTVSETDQSTFVREVRDLW  
LESLTPKSPASVSDSFGLGHSILATKFTLKKLQVDPILFTIFRPTIKAFAP  
EDIRKSGSSGSGQEVENVNTANAEAKLVETLPSSEYREYFEPNSAEKTTIN  
VEVTGVTGLSGYILADLLKSRPNYSFYFAHVRADDEAFARLQKAGITGTWNE  
KEASNIKVGLDLSKQFGLSDERKMLANVDLIIHNGALVHVYPAKLRDNPVLS  
TINVMALAAVGPKEFFDFVSTLCTCTEYFVNLKNSGKPGCIIESDGLMNSAGL  
TGCGYSKWAAYEIIIPRAGEFGLRGCIVRPGYVTGASVNGSNDFDGLLRLKGSVOL  
GKIPDIENVMVVPDVHVARVATSLNPKNELEVAQVTHPRILFKDYLYLHAY  
GDVETESYKWKKSLEASVDRNEENALYPLHMLVDNLNPESTKAPELDRNVAFL  
KDDAWIGVDWSNGTGVTFEEVGIYIAFLNKGVLFPPTTHNDKLPPLPSIELTQAQ:SL  
VAGAGAGSSAAA  
complement(6709..7185)  
/standard\_name="f(-) origin"  
7186..7496  
/gene="LacZ"  
/gene="LacZ"

RESULT 5 SYNPR8V 9655 bp DNA circular SYN 04-SEP-1997

LOCUS Cloning vector pGR8, complete sequence.

DEFINITION A001531

ACCESSION g2340041

KEYWORDS

ORGANISM

Cloning vector pGR8.  
Cloning vector pGR8  
artificial sequence; cloning vectors

QY 4453 GTAGTAACATACGATACGGAGGGCTTACCATCTGGCCCGAGTGTGCAATGATACCGG 4512  
D 4453 GTAGTAACATACGATACGGAGGGCTTACCATCTGGCCCGAGTGTGCAATGATACCGG 4512  
Db 3442 AGACCCACGCTACCGGCTCCAGATTATACGCAATAAACACGACGCGGAGGCGCGA 3501  
D 3442 AGACCCACGCTACCGGCTCCAGATTATACGCAATAAACACGACGCGGAGGCGCGA 3501  
QY 4513 AGACCCACGCTACCGGCTCCAGATTATACGCAATAAACACGACGCGGAGGCGCGA 4572  
D 4513 AGACCCACGCTACCGGCTCCAGATTATACGCAATAAACACGACGCGGAGGCGCGA 4572  
Db 3502 GCGCAGAGTGGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTATTTGTCGCGGA 3561  
D 3502 GCGCAGAGTGGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTATTTGTCGCGGA 3561  
QY 4573 GCGCAGAGTGGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTATTTGTCGCGGA 4632  
D 4573 GCGCAGAGTGGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTATTTGTCGCGGA 4632  
Db 3562 AGCTAGAGTAAGTAGTTCGCGCAGTTAATAGTTGCGCAACGTTGTCGCAATTGCTACAG 3621  
D 3562 AGCTAGAGTAAGTAGTTCGCGCAGTTAATAGTTGCGCAACGTTGTCGCAATTGCTACAG 3621  
QY 4633 AGCTAGAGTAAGTAGTTCGCGCAGTTAATAGTTGCGCAACGTTGTCGCAATTGCTACAG 4692  
D 4633 AGCTAGAGTAAGTAGTTCGCGCAGTTAATAGTTGCGCAACGTTGTCGCAATTGCTACAG 4692  
Db 3622 CATCTGCTGTCACGCTCGTGTGTTGATGCTGTCATTCAGTCCGGTTCCTACGATC 3681  
D 3622 CATCTGCTGTCACGCTCGTGTGTTGATGCTGTCATTCAGTCCGGTTCCTACGATC 3681  
QY 4693 CATCTGCTGTCACGCTCGTGTGTTGATGCTGTCATTCAGTCCGGTTCCTACGATC 4752  
D 4693 CATCTGCTGTCACGCTCGTGTGTTGATGCTGTCATTCAGTCCGGTTCCTACGATC 4752  
Db 3682 AAGSCGAGTTACATATCCCATGTTGTGCAAAAAGCGGTTAGCTCTCGGCTCTCC 3741  
D 3682 AAGSCGAGTTACATATCCCATGTTGTGCAAAAAGCGGTTAGCTCTCGGCTCTCC 3741  
QY 4753 AAGSCGAGTTACATATCCCATGTTGTGCAAAAAGCGGTTAGCTCTCGGCTCTCC 4812  
D 4753 AAGSCGAGTTACATATCCCATGTTGTGCAAAAAGCGGTTAGCTCTCGGCTCTCC 4812  
Db 3742 GATCTGTTGCAAGTAAGTTGGCGCAGTGTATATCACTCATGCTTATGCCACACTGCA 3801  
D 3742 GATCTGTTGCAAGTAAGTTGGCGCAGTGTATATCACTCATGCTTATGCCACACTGCA 3801  
QY 4813 GATCTGTTGCAAGTAAGTTGGCGCAGTGTATATCACTCATGCTTATGCCACACTGCA 4872  
D 4813 GATCTGTTGCAAGTAAGTTGGCGCAGTGTATATCACTCATGCTTATGCCACACTGCA 4872  
Db 3802 TAATCTCTTACTGTCATCCATCGTAAAGTGTCTTCTGACTCGTGGTAAACGTTCTTC 3861  
D 3802 TAATCTCTTACTGTCATCCATCGTAAAGTGTCTTCTGACTCGTGGTAAACGTTCTTC 3861  
QY 4873 TAATCTCTTACTGTCATCCATCGTAAAGTGTCTTCTGACTCGTGGTAAACGTTCTTC 4932  
D 4873 TAATCTCTTACTGTCATCCATCGTAAAGTGTCTTCTGACTCGTGGTAAACGTTCTTC 4932  
Db 3862 CAAGTCATCTGAGATAGTATGCGGCGACGAGTTGCTTCTGCGGCGGTCATACG 3921  
D 3862 CAAGTCATCTGAGATAGTATGCGGCGACGAGTTGCTTCTGCGGCGGTCATACG 3921  
QY 4933 CAAGTCATCTGAGATAGTATGCGGCGACGAGTTGCTTCTGCGGCGGTCATACG 4992  
D 4933 CAAGTCATCTGAGATAGTATGCGGCGACGAGTTGCTTCTGCGGCGGTCATACG 4992  
Db 3922 GGATAATACCGCGGCACATAGCAGACTTTTAAAGTGTCTATGATGGAACGTTCTTC 3981  
D 3922 GGATAATACCGCGGCACATAGCAGACTTTTAAAGTGTCTATGATGGAACGTTCTTC 3981  
QY 4993 GGATAATACCGCGGCACATAGCAGACTTTTAAAGTGTCTATGATGGAACGTTCTTC 5052  
D 4993 GGATAATACCGCGGCACATAGCAGACTTTTAAAGTGTCTATGATGGAACGTTCTTC 5052  
Db 3982 GGGGCGAAACTCTCAAGGATCTTACCGGTTGATGATCCAGTTCGATGTAACCCACTCG 4041  
D 3982 GGGGCGAAACTCTCAAGGATCTTACCGGTTGATGATCCAGTTCGATGTAACCCACTCG 4041  
QY 5053 GGGGCGAAACTCTCAAGGATCTTACCGGTTGATGATCCAGTTCGATGTAACCCACTCG 5112  
D 5053 GGGGCGAAACTCTCAAGGATCTTACCGGTTGATGATCCAGTTCGATGTAACCCACTCG 5112  
Db 4042 TCCACCCAACTGATCTTCAGCATCTTTTACTTTTCCACGCGTTCTTGGTGAGCAAAAC 4101  
D 4042 TCCACCCAACTGATCTTCAGCATCTTTTACTTTTCCACGCGTTCTTGGTGAGCAAAAC 4101  
QY 5113 TCCACCCAACTGATCTTCAGCATCTTTTACTTTTCCACGCGTTCTTGGTGAGCAAAAC 5172  
D 5113 TCCACCCAACTGATCTTCAGCATCTTTTACTTTTCCACGCGTTCTTGGTGAGCAAAAC 5172  
Db 4102 AGCAAGGCAAAATCCGCAAAAGGGAATAAGGCGCACACGGAATGTTGAATCTCAT 4161  
D 4102 AGCAAGGCAAAATCCGCAAAAGGGAATAAGGCGCACACGGAATGTTGAATCTCAT 4161  
QY 5173 AGCAAGGCAAAATCCGCAAAAGGGAATAAGGCGCACACGGAATGTTGAATCTCAT 5232  
D 5173 AGCAAGGCAAAATCCGCAAAAGGGAATAAGGCGCACACGGAATGTTGAATCTCAT 5232  
Db 4162 ACTCTCTCTTTTCAATATTTGAAGCATTTATCAGGGTTATTGCTCATGAGCGGATA 4221  
D 4162 ACTCTCTCTTTTCAATATTTGAAGCATTTATCAGGGTTATTGCTCATGAGCGGATA 4221  
QY 5233 ACTCTCTCTTTTCAATATTTGAAGCATTTATCAGGGTTATTGCTCATGAGCGGATA 5292  
D 5233 ACTCTCTCTTTTCAATATTTGAAGCATTTATCAGGGTTATTGCTCATGAGCGGATA 5292  
Db 4222 CATATTGATGATTTAGAAAATAAACAATAGGGGTTCCGGGCACATTTCCCGGAAA 4281  
D 4222 CATATTGATGATTTAGAAAATAAACAATAGGGGTTCCGGGCACATTTCCCGGAAA 4281  
QY 5293 CATATTGATGATTTAGAAAATAAACAATAGGGGTTCCGGGCACATTTCCCGGAAA 5352  
D 5293 CATATTGATGATTTAGAAAATAAACAATAGGGGTTCCGGGCACATTTCCCGGAAA 5352  
Db 4282 AGTGGCAC 4289  
D 4282 AGTGGCAC 4289  
QY 5353 AGTGGCAC 5360  
D 5353 AGTGGCAC 5360



US-09-020-716-3.rge

Wed Mar 17 09:43:33 1999

QY	4873	TAATCTCTTACTGTCATCCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTCAAC	4932		Cp	5240	GGAAGAGTATGAGTATTAACATTTCCCGTGTGGCCCTTATTCCTTTTTCGGGCATTTT	5181	
Db	9220	CAAGTCAATCTGAGATAGTGTATGCGGAGCCAGATGCTCTTGGCCGGCGTCAATACG	9279		Db	1453	GCCTTCCTGTTTTTGTCTACCCAGAAACGCTGTGAAAGTAAAGATGCTGAAGATCAGT	1512	
QY	4933	CAATCTATTTGAGATAGTATGCGGAGCCGAGATGCTCTTGGCCGGCGTCAATACG	4992		Cp	5180	GCCTTCCTGTTTTTGTCTACCCAGAAACGCTGTGAAAGTAAAGATGCTGAAGATCAGT	5201	
Db	9280	CGAATATACCGCGCCACATAGCAGAACCTTTAAAGTGTCTATCATTTGGAAACGTTCTTC	9339		Db	1513	TGGGTGCACAGTGGGTGTACATCGAACTGGATCTCAACAGCGTAAAGATCTTGGAGATT	1572	
QY	4993	GGATATACCGCGCCACATAGCAGAACCTTTAAAGTGTCTATCATTTGGAAACGTTCTTC	5052		Cp	5120	TGGGTGCACAGTGGGTGTACATCGAACTGGATCTCAACAGCGTAAAGATCTTGGAGATT	5061	
Db	9340	GGGCGGAAACCTCTCAAGGATCTTACCGTGTGTGAGATCCAGTTCGATGAACCACTCG	9399		Db	1573	TTCCGCGGAGAGACGTTTTTCCAAATGATGAGCACITTTAAAGTCTGTGTATGTCTGCGG	1632	
QY	5053	GGGCGGAAACCTCTCAAGGATCTTACCGTGTGTGAGATCCAGTTCGATGAACCACTCG	5112		Cp	5060	TTCCGCGGAGAGACGTTTTTCCAAATGATGAGCACITTTAAAGTCTGTGTATGTCTGCGG	5001	
Db	9400	TGAATCAAGTCTCTCAAGGATCTTACCGTGTGTGAGATCCAGTTCGATGAACCACTCG	9459		Db	1633	TATTAATCCGCTATTGAGCGCGGCGGAGAGCACTCGTGTGCGGCATACATATCTCAGA	1692	
QY	5113	TGAATCAAGTCTCTCAAGGATCTTACCGTGTGTGAGATCCAGTTCGATGAACCACTCG	5172		Cp	5000	TATTAATCCGCTATTGAGCGCGGCGGAGAGCACTCGTGTGCGGCATACATATCTCAGA	4941	
Db	9460	AGGAAGGCAAAATGCGGCAAAAGCGAATAGGGGCGACAGGAAATGTTGAATACTCAT	9519		Db	1693	ATGACTTGGTGTGAGTACTCACCAGTACAGAAAGAGCATTTACGATGACATTAAGTAA	1752	
QY	5173	AGGAAGGCAAAATGCGGCAAAAGCGAATAGGGGCGACAGGAAATGTTGAATACTCAT	5232		Cp	4940	ATGACTTGGTGTGAGTACTCACCAGTACAGAAAGAGCATTTACGATGACATTAAGTAA	4881	
Db	9520	ATGCTTCTTTTCAATATATTGAAGCATTTATCAGGTTATTGTCTCATGAGCGGATA	9579		Db	1753	GAGAATTTATGAGTGTCTGCTATACCATGAGTGTATACACTGCGGCGCAACTTCTGA	1812	
QY	5233	ATGCTTCTTTTCAATATATTGAAGCATTTATCAGGTTATTGTCTCATGAGCGGATA	5292		Cp	4880	GAGAATTTATGAGTGTCTGCTATACCATGAGTGTATACACTGCGGCGCAACTTCTGA	4821	
Db	9580	CATATTTGAATGTTATGAAATAAACAATAGGGGTTCCGCGCACATTTCCCGGAAA	9639		Db	1813	CAAGCATCGGAGGCGGAGGAGCTAACCGCTTTTTTGCACACATGCGGATCATGTAA	1872	
QY	5293	CATATTTGAATGTTATGAAATAAACAATAGGGGTTCCGCGCACATTTCCCGGAAA	5352		Cp	4820	CAAGCATCGGAGGCGGAGGAGCTAACCGCTTTTTTGCACACATGCGGATCATGTAA	4761	
Db	9640	AGTGGCAC 9647			Db	1873	CTCGCCTTGTATGCTTGGGAAACCGGAGCTGAATGAGCCATACCAAGAGCGAGTGACA	1932	
QY	5353	AGTGGCAC 5360			Cp	4760	CTCGCCTTGTATGCTTGGGAAACCGGAGCTGAATGAGCCATACCAAGAGCGAGTGACA	4701	
RESULT	5	AFU25397	9250 bp	DNA	SYN	17-DEC-1997			
LOCUS		AFU25397							
DEFINITION		Expression vector pesp-3, complete sequence.							
ACCESSION		AF025397							
NID		q2689634							
KEYWORDS		Expression vector pesp-3.							
SOURCE		Expression vector pesp-3.							
ORGANISM		artificial sequence; expression vectors.							
REFERENCE		1 (bases 1 to 9750)							
AUTHORS		Lu, Q.							
TITLE		Direct Submission							
JOURNAL		Submitted (16-SEP-1997) Marketing Analysis, Stratagene, 11011 North							
COMMENT		Torrey Pines Road, La Jolla, CA 92037, USA							
FEATURES		On Dec 17, 1997 this sequence version replaced gi.2623066							
		Location/Qualifiers							
		1..9750							
		/organism="Expression vector pesp-3"							
		/specific_host="Schizosaccharomyces pombe"							
		/db_xref="taxon:57775"							
		/plasmid="PESP-3"							
BASE COUNT		2919 a 1959 c 1983 g 2889 t							
ORIGIN									
		Query Match 41.6%; Score 2228; DB 32; Length 9750;							
		Best Local Similarity 100.0%; Pred No. 0.00e+00;							
		Matches 2228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
Db	1273	GUGGCACTTTTCGGGAAATGTCGCGGAAACGCTTATTTGTTATTTTCTTAATACATT	1332		Db	4281	AGATTGATTTAAACCTTCATTTTAAATTTAAAGAGTCTAGGTGAAGATGCTTTTGATA	4281	
Cp	5360	GUGGCACTTTTCGGGAAATGTCGCGGAAACGCTTATTTGTTATTTTCTTAATACATT	5301		Db	4340	AGATTGATTTAAACCTTCATTTTAAATTTAAAGAGTCTAGGTGAAGATGCTTTTGATA	4281	
Db	1333	CAATATGATTCGGCTCATGACACATACCTGATAAATGCTTCAATATATGAAAA	1392		Db	2233	TAGTTCCTCTACTGATTAAAGCATTTGTTAACTGTCAGACCAAGTTTACTCATATATACIT	2292	
Cp	5300	CAATATGATTCGGCTCATGACACATACCTGATAAATGCTTCAATATATGAAAA	5241		Cp	4400	TAGTTCCTCTACTGATTAAAGCATTTGTTAACTGTCAGACCAAGTTTACTCATATATACIT	4341	
Db	1393	GGAAGAGTATGAGTATTAACATTTCCGTTGTGGCCCTTATTCCTTTTTCGGGCATTTT	1452		Db	2293	AGATTGATTTAAACCTTCATTTTAAATTTAAAGAGTCTAGGTGAAGATGCTTTTGATA	2352	



5180 GCCTTCCTGTTTTTGTCTACCCCAAGAACCGTGGTGAAGCTAAAGATGCTGAAGATCAGT 5121

1513 TGGGTGGACAGATGGGTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGGAGGTT 1572

5120 TGGGTGGACAGATGGGTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGGAGGTT 5061

1573 TTCGCCGGGAAGACGTTTTCCATGATGAGCACTTTTAAAGTCTCTGCTATGTGCGCGGG 1632

5059 TTAGCGGGGAAATAAAGCTTTTCCAAATGATGAGCACTTTTAAAGTCTCTGCTATGTGCGCGGG 5001

1533 TATTAATCGCGTATTAGACGCCGGGAAAGCAACTCGGTCCGCCGATACACTATTCTCAGA 1692

5000 TATTAATCGCGTATTAGACGCCGGGAAAGCAACTCGGTCCGCCGATACACTATTCTCAGA 4941

1693 ATGACATTGTTTGGTACTCACCAGTACAGAAAAGCATCTTACGGATGGCATGACAGATAA 1752

4940 ATGACATTGTTTGGTACTCACCAGTACAGAAAAGCATCTTACGGATGGCATGACAGATAA 4881

1753 GAGAAATTACGAGTGGTGGCCATAAACCATGATGATAAACACTTGGCGCCAACTTACTTCTGA 1812

4890 GAGAAATTACGAGTGGTGGCCATAAACCATGATGATAAACACTTGGCGCCAACTTACTTCTGA 4821

1813 CAACGATCGGAGGACCGAAGGACTTAACCGCTTTTTTGCACAACATGGGGCATCATGTAA 1872

4820 CAACGATCGGAGGACCGAAGGACTTAACCGCTTTTTTGCACAACATGGGGCATCATGTAA 4761

1873 CTCGGCTTGATCGTTTGGGAACCGGAGCTGAAAGAACCATTACCAACGACGAGGTGACA 1932

4760 CTCGGCTTGATCGTTTGGGAACCGGAGCTGAAAGAACCATTACCAACGACGAGGTGACA 4701

1933 GCACAGATGCTTGTAGCAATCGCACAACTTTTGGCGCAAACTATTAACTGGCGCAACTACTTA 1992

4700 GCACAGATGCTTGTAGCAATCGCACAACTTTTGGCGCAAACTATTAACTGGCGCAACTACTTA 4641

1993 CTCACGCTCCGGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGAGCACAC 2052

4640 CTCACGCTCCGGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGAGCACAC 4581

2053 TTCGCGCTTCGCGCTTCGAGTGGCTGGTATTGCTGTGATAAATCTGAGCGGTGAGC 2112

4580 TTCGCGCTTCGCGCTTCGAGTGGCTGGTATTGCTGTGATAAATCTGAGCGGTGAGC 4521

2113 GTGGATCTCGCGGTATCATTCGAGAACTTGGGCGAGATGGTAAGCCCTCCCGTATCGTAG 2172

4520 GTGGATCTCGCGGTATCATTCGAGAACTTGGGCGAGATGGTAAGCCCTCCCGTATCGTAG 4461

2173 TTATCTACAGACGGGGAGTCAGGCAACTATGGATCAACGAAATAGACAGATCGCTGAGA 2232

4460 TTATCTACAGACGGGGAGTCAGGCAACTATGGATCAACGAAATAGACAGATCGCTGAGA 4401

2233 TAAATGCTTCACTGATTAAAGCATTGGTAAGTTCAGCAAGTTCAGCTATATATCTTT 2292

4400 TAAATGCTTCACTGATTAAAGCATTGGTAAGTTCAGCAAGTTCAGCTATATATCTTT 4341

2293 AGATTGATTTAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCCTTTTGAATA 2352

4340 AGATTGATTTAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCCTTTTGAATA 4281

2353 ATCTCATGACCAAAATCCCTTAAAGTTCAGCTTTCGCTTCCACTGAGCGTCAGACCCCGTAG 2412

4280 ATCTCATGACCAAAATCCCTTAAAGTTCAGCTTTCGCTTCCACTGAGCGTCAGACCCCGTAG 4221

2413 AAAAGATCAAGGATCTCTCTGAGATCCCTTTTTTCTGCGCGTAATCTGCTGCTGCATAA 2472

4220 AAAAGATCAAGGATCTCTCTGAGATCCCTTTTTTCTGCGCGTAATCTGCTGCTGCATAA 4161

2473 CAAAAAAACCAAGGATCTTACGAGCGGTGGTTTGGTTCCGGGATCAGAGCTACCAACTCTTT 2532

4160 CAAAAAAACCAAGGATCTTACGAGCGGTGGTTTGGTTCCGGGATCAGAGCTACCAACTCTTT 4101

2533 TCTGAGGTAAGTCTGCTTTCAGCAGAGCGAGATACCAAAATCTGCTCTTCTAGTGTAGC 2592

4100 TCTGAGGTAAGTCTGCTTTCAGCAGAGCGAGATACCAAAATCTGCTCTTCTAGTGTAGC 4041

Db	2593	CGTAGTTAGGCACACACTTCAAGAACTCTGTAGCAGCGGCTACATACATCCTGCTGCTAA	2652
Cp	4040	CGTAGTTAGGCACACACTTCAAGAACTCTGTAGCAGCGGCTACATACATCCTGCTGCTAA	3981
Db	2653	TGCTGTACCACTGGCTGCTGCCAGTAGTGGCATTAAGTCGTCTTACCGGGTTGGACTCAA	2712
Cp	3980	TGCTGTACCACTGGCTGCTGCCAGTAGTGGCATTAAGTCGTCTTACCGGGTTGGACTCAA	3921
Db	2713	GACGATAGTTACCGGATAGCGCAGAGGGGTGGGCTGAAGCGGGGGTTGCTGCAACACAG	2772
Cp	3920	GACGATAGTTACCGGATAGCGCAGAGGGGTGGGCTGAAGCGGGGGTTGCTGCAACACAG	3861
Db	2773	CCAGCTTGGAGCGAAGCAGCTACACCGAACTTACATACCTACAGCGTGACTATGAGAAA	2832
Cp	3860	CCAGCTTGGAGCGAAGCAGCTACACCGAACTTACATACCTACAGCGTGACTATGAGAAA	3801
Db	2833	GCGCCAGGCTCCCGAAGGAGAAAGGGGACAGGTATCCGGTAAGCGGCAGGGTCGAA	2892
Cp	3800	GCGCCAGGCTCCCGAAGGAGAAAGGGGACAGGTATCCGGTAAGCGGCAGGGTCGAA	3741
Db	2893	CAGGAGAGCGCAGAGGAGCTTCCAGGGGAAAGCGCTGTATCTTATAGTCTGCTG	2952
Cp	3740	CAGGAGAGCGCAGAGGAGCTTCCAGGGGAAAGCGCTGTATCTTATAGTCTGCTG	3681
Db	2953	GGTTGGCCACTCTGACITGAGCGTCGATTTTGTGATGCTGCTGCTGCTGCTGCTGCTG	3012
Cp	3680	GGTTGGCCACTCTGACITGAGCGTCGATTTTGTGATGCTGCTGCTGCTGCTGCTGCTG	3421
Db	3013	TATGAAAGAGCGCAGAGCGGCGCTTTTACGGTTCCTGGCCCTTTTGTGCTGGCCTTTG	3072
Cp	3620	TATGAAAGAGCGCAGAGCGGCGCTTTTACGGTTCCTGGCCCTTTTGTGCTGGCCTTTG	3561
Db	3073	CTCACATGTTCTTCTCGTGTATCCCTGATTCGTGCGATACCGGTATACCGGCTTG	3132
Cp	3560	CTCACATGTTCTTCTCGTGTATCCCTGATTCGTGCGATACCGGTATACCGGCTTG	3501
Db	3133	AGTAGCTGATACCGCTGGCCAGCGCAAGCAGCGAGCGACCGGCTAGTGAAGGAGG	3192
Cp	3500	AGTAGCTGATACCGCTGGCCAGCGCAAGCAGCGAGCGACCGGCTAGTGAAGGAGG	3441
Db	3193	AAGGGGAAGAGCGCTTAATACGCAAAACCGCTCTCCCGCGCGTTGGAGCAATCATTAAT	3252
Cp	3440	AAGGGGAAGAGCGCTTAATACGCAAAACCGCTCTCCCGCGCGTTGGAGCAATCATTAAT	3381
Db	3253	GCAGCTGGCAGCAGGTTTCCGAGCTGAAAGCGGCGAGTGAAGGCAAGCAATTAATG	3312
Cp	3380	GCAGCTGGCAGCAGGTTTCCGAGCTGAAAGCGGCGAGTGAAGGCAAGCAATTAATG	3321
Db	3313	TGAGTTAGCTCACTCATTTAGGCAACCGAGCTTTACATTTATGCTTCGGCTGCTATGI	3372
Cp	3320	TGAGTTAGCTCACTCATTTAGGCAACCGAGCTTTACATTTATGCTTCGGCTGCTATGI	3261
Db	3373	TGTTGGAAATTTGAGCGGATACAAATTTACACAGGAACAGCTATGACCATGATACG	3432
Cp	3260	TGTTGGAAATTTGAGCGGATACAAATTTACACAGGAACAGCTATGACCATGATACG	3201
Db	3433	CCAAGCGGCAATTAACCCCTCACTTAAGGGAACAAAGCTGGGTACGAGCGGCGGCTCG	3492
Cp	3200	CCAAGCGGCAATTAACCCCTCACTTAAGGGAACAAAGCTGGGTACGAGCGGCGGCTCG	3141
Db	3493	AGGTCGAC	3500
Cp	3140	AGGTCGAC	3133

8	EVU67875	10765 bp	DNA	circular	SYN	16-DEC-1987
RESULT						
LOCUS						
DEFINITION	pSP-1 yeast expression vector, complete sequence.					
ACCESSION	U67875					
VERSION	g2689262					
NID						
KEYWORDS	.					
SOURCE	Expression vector pSP-1.					



Wed Mar 17 09:43:33 1999

Cp 4100 CAAAAAACACAGCGCTACGAGCGGTGGTTGTTTTCGCGGATCAAGAGCTACCAACICTTT 4101  
 Db 2523 TTCCGACGTAAGTGGCTTCAGCAGAGCGCAGATACCAATACTCTCTCTAGTGTAGC 2592  
 Cp 4100 TTCCGACGTAAGTGGCTTCAGCAGAGCGCAGATACCAATACTCTCTCTAGTGTAGC 4041  
 Db 2593 CTAAGTGTAGTGTACCTTCAGCAAGCTCTGTAGCAGCGCTACATACCTCTCTGCTAA 2652  
 Cp 4040 CGTAGTGTAGCGCACCTTCAGCAAGCTCTGTAGCAGCGCTACATACCTCTCTGCTAA 3981  
 Db 2653 TCTGTGTAGCAGTGGCTTCAGCAGTGGCGATAGTCTGTGTACCGGTTGGACTCAA 2712  
 Cp 3981 TCTGTGTAGCAGTGGCTTCAGCAGTGGCGATAGTCTGTGTACCGGTTGGACTCAA 3921  
 Db 2713 GACGATAGTGTACGATTAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGTCACACAGC 2772  
 Cp 3921 GACGATAGTGTACGATTAAGCGGAGCGGTCGGGCTGAAAGCGGGGTTGTCACACAGC 3861  
 Db 2773 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 2832  
 Cp 3861 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3801  
 Db 2833 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 2892  
 Cp 3801 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3741  
 Db 2892 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 2952  
 Cp 3741 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3681  
 Db 2953 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3012  
 Cp 3681 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3621  
 Db 3012 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3072  
 Cp 3621 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3561  
 Db 3072 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3132  
 Cp 3561 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3501  
 Db 3132 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3192  
 Cp 3501 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3441  
 Db 3192 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3252  
 Cp 3441 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3381  
 Db 3252 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3312  
 Cp 3381 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3321  
 Db 3312 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3372  
 Cp 3321 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3261  
 Db 3372 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3432  
 Cp 3261 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3201  
 Db 3432 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3492  
 Cp 3201 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3141  
 Db 3492 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3500  
 Cp 3141 CAGCTTGTAGGAGGAGCAGCTACAGCGAGCTGAGATACCTACAGCGTGTAGCTATGAGAAA 3133

RESULT 9 SYNBLKSPV 2964 bp DNA circular SYN 04-JUN-1993  
 LOCUS

DEFINITION BlueScribe KS Plus cloning vector.  
 ACCESSION L08785  
 NID 9310729  
 KEYWORDS Synthetic construct DNA.  
 ORGANISM artificial sequence.  
 SOURCE 1 (bases 1 to 2964)  
 REFERENCE Gilbert, W.  
 TITLE Obtained from Vecbase 3.0  
 JOURNAL Unpublished (1991)  
 COMMENT These data and their annotation were supplied to GenBank by Will Gilbert under the auspices of the GenBank Curator Program.  
 ENTRY BLUEKSP  
 TITLE BlueScribe KS Plus - Cloning vector  
 DATE 28-JAN-1987  
 #sequence 02-FEB-1987  
 #sequence 04-MAR-1987  
 #sequence 03-APR-1987  
 ACCESSION V00078  
 SOURCE artificial  
 REFERENCE  
 #number 1  
 #authors Fernandez J.M., Short J.M., Kershaw M., Hare W., Scribe J  
 #journal Gene (1987) in press  
 #citation Sequence data from Stratagene  
 #comment sequence correction according to Stratagene COMMENT  
 #obtained from Stratagene on floppy disc.  
 Revised 02-FEB-1987 by F. Pfeiffer;  
 14/09/10 'AT' to 'TA' to match revised sequence of pBR322  
 Revised 4-MAR-1987 to match sequence of pUC19 on request  
 of Stratagene  
 Polylinker region revised 03-APR-1987 according to Stratagene  
 COMMENT  
 The stand shown corresponds to pUC19c.  
 As in the published sequence of pUC19c, The M13mp18 lacZ region  
 is on the complementary strand.  
 This vector contains the fl origin so that the plus strand  
 can be obtained upon fl superinfection.  
 KEYWORDS  
 CROSSREFERENCE  
 #parent  
 Vecbase(3):BlueM13p  
 #parent  
 Vecbase(3):pUC19, VecSource(3):PromT7, VecSource(3):PromT3.  
 VecSource(3):bGalkS, GenBank(50):PFI  
 #brother  
 VecBase(3):BlueKSm, VecBase(3):BlueSkp  
 PARENT  
 Features of BlueKsp (2964 bp)  
 residue source  
 3- 458 5943-5488 (c) phage fl  
 460- 624 236- 400 pUC19c  
 626- 645 1- 20 T7 promoter  
 653- 760 108- 1 (c) BlueKs-polylinker  
 772- 791 20- 1 (c) T3 promoter  
 795-2964 448-2617 pUC19c  
 Conflict (cfl) and Mutations (mut): none  
 PARENT  
 Features of BlueKsp (2964 bp)  
 residue source  
 3- 458 5943-5488 (c) phage fl  
 460- 624 236- 400 pUC19c  
 626- 645 1- 20 T7 promoter  
 653- 760 108- 1 (c) BlueKs-polylinker  
 772- 791 20- 1 (c) T3 promoter  
 795-1031 237- 1 (c) pUC19  
 1032-2964 268- 754 (c) pUC19  
 Conflict (cfl) and Mutations (mut): none















[illegible]



Wed Mar 17 09:43:33 1999

**KEYWORDS**  
**SOURCE** Cloning vector pFUN  
**ORGANISM** Cloning vector pFUN  
**REFERENCE** Artificial sequence: cloning vectors.  
**AUTHORS** 1 (bases 1 to 8072)  
**TITLE** Poquet, I., Ehrlich, S.D. and Gruss, A.  
**JOURNAL** An export-specific reporter designed for gram-positive bacteria:  
**MEDLINE** application to lactococcus lactis  
**REFERENCE** J. Bacteriol. 180 (7), 1904-1912 (1998)  
**AUTHORS** 2 (bases 1 to 8072)  
**TITLE** Poquet, I. and Gruss, A.  
**JOURNAL** Direct Submission  
**REFERENCE** Submitted (15-DEC-1997) Laboratoire de Genetique Appliquee-URLGA,  
**AUTHORS** Institut National de la Recherche Agronomique, CRJ, Jouy en Josas  
**TITLE** 78352, France  
**FEATURES** Location/Qualifiers  
**source** 1..8072  
**misc\_feature** /organism="Cloning vector pFUN"  
**misc\_feature** /db\_xref="taxon:74538"  
**terminator** /note="plasmid vector that creates translational fusions  
with a nuclelease reporter without a signal peptide"  
**misc\_feature** 1..4580  
**gene** /note="pL252 moiety: similar to sequence deposited under  
GenBank Accession Number AF039139"  
**misc\_feature** 4581..7472  
**terminator** /note="similar to BlueScript"  
**misc\_feature** 7473..7507  
**misc\_feature** /note="trpA terminator"  
**gene** 7512..7547  
**misc\_feature** /note="multiple cloning site"  
**gene** <7543..8010  
**CDS** /note="truncated nuc gene potentially encoding deltaSP-Nuc  
ORF"  
**misc\_feature** /gene="delta-nuc"  
**misc\_feature** <7543..8010  
**misc\_feature** /gene="delta-nuc"  
**misc\_feature** /note="nuclelease from Staphylococcus aureus devoid of its  
signal peptide and the thirteen amino acids at the  
N-terminus of mature protein"  
**misc\_feature** /codon\_start=1  
**misc\_feature** /transl\_table=1  
**misc\_feature** /product="deltaSP-Nuc"  
**misc\_feature** /db\_xref="PIB:G3043925"  
**misc\_feature** /translation="DPTVYSATSKLHKPEATLLKIDGDTVKLMYKQPMTERLLL  
VDTPETKPKRGVKGPEASATKKWENAKKIEVEDKQRTDKYGRGLAYVADG  
KWNVEALVPGGLAKVAYVYKPNNTKSHOHLKSEAQKKEKLINSEDNADSQ"  
**BASE COUNT** 2717 a 1474 c 1694 g 2187 t  
**ORIGIN**  
**Query Match** 41.5%; Score 2224; Pos 32; Length 8072;  
**Best Local Similarity** 44.0%; Pred No. 0.00e+00;  
**Matches** 2225; Conservative 0; Mismatches 2; Indels 0; Gaps 0.  
**Db** 5240 GTGCGACCTTTTCGGGGAATGTGCGGGAACCCCTATTGTTTATTTCTAATACATT 5299  
**Cp** 5360 GTGCGACCTTTTCGGGGAATGTGCGGGAACCCCTATTGTTTATTTCTAATACATT 5301  
**Db** 5360 CAATATGATGCTGCTCATGAGAAATAAATACCTGATATAATGCTTCAATAATATGAAAAA 5359  
**Cp** 5360 CAATATGATGCTGCTCATGAGAAATAAATACCTGATATAATGCTTCAATAATATGAAAAA 5241  
**Db** 5360 GGAAGAGTATGAGTATTAACATTTTCGGTGGCCCTTTATCCCTTTTTCGGGCATTTT 5419  
**Cp** 5240 GGAAGAGTATGAGTATTAACATTTTCGGTGGCCCTTTATCCCTTTTTCGGGCATTTT 5181  
**Db** 5420 GCTTTCCTGTTTTCGGTGGCCCTTTATCCCTTTTTCGGGCATTTT 5419  
**Cp** 5180 GCTTTCCTGTTTTCGGTGGCCCTTTATCCCTTTTTCGGGCATTTT 5121  
**Db** 5480 TGGTTCCTGTTTTCGGTGGCCCTTTATCCCTTTTTCGGGCATTTT 5539  
**Cp** 5120 TGGTTCCTGTTTTCGGTGGCCCTTTATCCCTTTTTCGGGCATTTT 5061

5540 TTGCGCCCGGAAGAGCTTTTCCAAATGATGAGCACTTTTAAAGTTTCTGCTATGTGGCGGG 5599  
5060 TTGCGCCCGGAAGAGCTTTTCCAAATGATGAGCACTTTTAAAGTTTCTGCTATGTGGCGGG 5001  
5600 TATTATCCGTTATGACCGCGGCAAGAGCACTTGGTCTGCGGCATACATATTTCTAGA 5659  
5000 TATTATCCGTTATGACCGCGGCAAGAGCACTTGGTCTGCGGCATACATATTTCTAGA 4941  
5660 ATGACTTGGTTGAGTACTCACCAGTTCACAGAAAGCATTTACGGATGGCATGACAGTAA 5719  
4940 ATGACTTGGTTGAGTACTCACCAGTTCACAGAAAGCATTTACGGATGGCATGACAGTAA 4881  
5720 GAGAAATATGCACTGCTGCATTAACCATGATGATGATGATGATGATGATGATGATGATG 5779  
4880 GAGAAATATGCACTGCTGCATTAACCATGATGATGATGATGATGATGATGATGATGATG 4821  
5780 CAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTCGACAACTATGCGGATCATGTAA 5839  
4820 CAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTCGACAACTATGCGGATCATGTAA 4761  
5840 CTCGCTTGTATGCTTGGGACCGGAGCTGAATGAAGGCAATATGCGGATCATGTAA 5899  
4760 CTCGCTTGTATGCTTGGGACCGGAGCTGAATGAAGGCAATATGCGGATCATGTAA 4701  
5900 CCAGATGCTGCTGAGCAATGGCAACACGTTGGGCAAACTATTAAGTTCGCAACTACTTA 5959  
4700 CCAGATGCTGCTGAGCAATGGCAACACGTTGGGCAAACTATTAAGTTCGCAACTACTTA 4641  
5960 CTCAGCTTCCCGGCAACAAATTAATAGACTGGATGAGGCGGATAAAGTTTCAGAGCACT 6019  
4640 CTCAGCTTCCCGGCAACAAATTAATAGACTGGATGAGGCGGATAAAGTTTCAGAGCACT 4581  
6020 TTCTCGCTCGCCCTTCGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 6079  
4580 TTCTCGCTCGCCCTTCGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 4521  
6080 GTGGGTCTCGGGGTATCATTCAGCACTGGGCGGCGAGATGTAAGCGCTCCCGTATGAGT 6139  
4520 GTGGGTCTCGGGGTATCATTCAGCACTGGGCGGCGAGATGTAAGCGCTCCCGTATGAGT 4461  
6140 TTATCTACACGACGGGGAGTCAGGCAACTATGATGAACCAATATAGACATGCGTGAGA 6199  
4460 TTATCTACACGACGGGGAGTCAGGCAACTATGATGAACCAATATAGACATGCGTGAGA 4401  
6200 TAGTCCCTCAGTATTAAGCATTCGTAACGTGTCAGACCAAGTTTACTCATATATACAT 6259  
4400 TAGTCCCTCAGTATTAAGCATTCGTAACGTGTCAGACCAAGTTTACTCATATATACAT 4341  
6260 AGATTGATTAAACCTTCATTTTAAATTTAAAGGATCTAGTGAAGATCGTTTTCATTA 6319  
4340 AGATTGATTAAACCTTCATTTTAAATTTAAAGGATCTAGTGAAGATCGTTTTCATTA 4281  
6320 ATCTCATGACCAAAATCCCTTAAGCTGAGTTTCTGTCACAGAGAGTTCAGACATGAGT 6379  
4280 ATCTCATGACCAAAATCCCTTAAGCTGAGTTTCTGTCACAGAGAGTTCAGACATGAGT 4221  
6380 AAAAGATCAAAAGGATCTTCTTGAGATCCCTTTTTCGGGCTATCTGCTGCTGCAAAA 6439  
4220 AAAAGATCAAAAGGATCTTCTTGAGATCCCTTTTTCGGGCTATCTGCTGCTGCAAAA 4161  
6440 CAAAAAACCACCTACAGCGGCTGGTTTGTTCGGGATCAAGAGCTACCAACTCTTT 6499  
4160 CAAAAAACCACCTACAGCGGCTGGTTTGTTCGGGATCAAGAGCTACCAACTCTTT 4101  
6500 TTCCGAAGTAACTGGGTTCAGCAGAGCGAGATACCAATACTGCTCCTTCTAGTGTAGC 6559  
4100 TTCCGAAGTAACTGGGTTCAGCAGAGCGAGATACCAATACTGCTCCTTCTAGTGTAGC 4041  
6560 CGTAGTTAGCGCACCTTCAAGACTCTGTAGACCGGCTACATACCTGGTGTCTGTAA 6619  
4040 CGTAGTTAGCGCACCTTCAAGACTCTGTAGACCGGCTACATACCTGGTGTCTGTAA 3981











Sequence of clone BP2 0.00e+00  
Plasmid pWRG3169 enco 0.00e+00  
Plasmid pM16-1. 0.00e+00  
Plasmid pM16. 0.00e+00  
Plasmid pVSV(-) lac 0.00e+00  
K.lactis/S. cerevisiae 0.00e+00  
Vector p37M1-10D. 0.00e+00  
Construct pGEM-HTR(Ka 0.00e+00  
Construct pGEM-HTR co 0.00e+00  
pDE110. 0.00e+00  
Plasmid pMP6-IL2. 0.00e+00  
Interleukin-2 expres 0.00e+00  
pVE108. 0.00e+00  
Plasmid pVE109 used i 0.00e+00  
pDE4. Plasmid DNA rep 0.00e+00  
Plasmid pWRG3196 enco 0.00e+00  
Sequence of pTE4. 0.00e+00  
Plasmid pCB51 encodin 0.00e+00  
Plasmid pCB50 encodin 0.00e+00  
Plasmid pCPM1.8 conta 0.00e+00  
Human cyclin D1-human 0.00e+00  
Shortened c1 gene in 0.00e+00  
pDE108. 0.00e+00  
pDE102/3 Apal constr 0.00e+00  
Plasmid pVE144 used i 0.00e+00  
IL-4.Y134D/Yq31 fusio 0.00e+00  
Plasmid pHCV-157 codi 0.00e+00  
pHCV167 sequence. 0.00e+00  
Plasmid pHCV-162 codi 0.00e+00  
Plasmid pTB72 encodin 0.00e+00  
Alphavirus-based euka 0.00e+00

14 2198 41.0 6526 7 C40281  
15 2198 41.0 7287 39 V02042  
16 2196 41.0 10930 20 Q81226  
17 2196 41.0 10950 20 Q81225  
18 2195 41.0 14311 24 Q38744  
19 2190 40.9 6824 6 Q39050  
20 2185 40.8 7228 6 Q50201  
21 2166 40.4 4118 35 T69189  
22 2166 40.4 4118 35 T69188  
23 2163 40.4 4883 4 Q25709  
24 2163 40.4 4883 8 Q51192  
25 2165 40.4 5585 14 Q87418  
26 2163 40.4 5585 29 T61430  
27 2163 40.4 5620 4 Q27489  
28 2163 40.4 5620 4 Q27489  
29 2163 40.4 5642 6 Q51193  
30 2167 40.4 6295 39 V02043  
31 2163 40.4 7050 7 Q40419  
32 2163 40.4 10288 31 T71322  
33 2163 40.4 13414 31 T71321  
34 2162 40.3 4454 30 T64622  
35 2162 40.3 4540 32 T69891  
36 2162 40.3 4824 20 T68975  
37 2162 40.3 5399 4 Q25706  
38 2162 40.3 6274 7 Q43812  
39 2162 40.3 6555 9 Q53874  
40 2162 40.3 6926 18 T12662  
41 2162 40.3 7106 8 Q47193  
42 2162 40.3 7106 15 Q97424  
43 2162 40.3 7298 8 Q47192  
44 2162 40.3 10443 31 T71319  
45 2162 40.3 16656 21 T30787

ALIGNMENTS

RESULT 1  
ID T49876 standard; DNA: 5178 BP.

AC T49876;  
DE 28-MAR-1997 (first entry)  
DE pTet-Splice.  
KW pTet-Splice; pTet-trak, tetracycline transactivator; promoter;  
KW gene expression; transgenic animal; animal model; drug screening;  
KW vector; ds.  
OS Synthetic.  
FH Key  
FI misc\_rna  
FI Location/Qualifiers  
FI complement (2591)  
FI /\*tag= a  
FI /\*note= "putative start site of trxn"  
FI complement (2616..2622)  
FI /\*tag= b  
FI tata\_signal  
FI W09640946-Al.  
PN 19-DEC-1996.  
PD 19-DEC-1996.  
PF 07-JUN-1996; U10109.  
PR 07-JUN-1995; US-474169.  
PA (UYVA ) UNIV YALE.  
PI Schatz DG;  
PI WPI: 97-077273/07.  
DR Nucleic acid encoding tetracycline transactivator fusion protein -  
PT provides rapid and reversible control of gene expression, e.g. for  
PT creating animal models for drug screening  
PS Example 1: Fig 9b-g, 82pp, English.  
CC Vector pTet-Splice (T49876) is utilised in the construction of  
CC autoregulatory vector Plasmid pTet-trak (T49877) Splice-PA was  
CC made by ligating the SV40 small T antigen intervening sequence and  
CC the SV40 early polyA sequence into pTetSII+. The XhoI-SalI fragment  
CC of pUC13-3 (contg. 7 copies of the tet operator upstream of minimal  
CC promoter rep) was cloned upstream of the splice/polyA sequence of  
CC pTet-Splice-PA to form pTet-Splice A modified tetracycline  
CC transactivator (ttak) gene was cloned into pTet-Splice to form pTet-  
CC ttak. The construct provides rapid, reversible control of gene  
CC expression in eukaryotic cells or transgenic animals, e.g. for  
CC heterologous protein, RNA or antisense sequence prodn., or for

\*\*\*\*\*  
Release 3.1A John F. Collins, Biocomputing Research Unit  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd  
\*\*\*\*\*  
MPSrch\_n n.a. - n.a. database search, using Smith-Waterman algorithm  
Run on Sun Mar 14 23:15:31 1999. MasPar time 757.38 Seconds  
962.844 Million cell updates/sec  
Tabular output not generated.  
\*\*\*\*\*  
Title: US-09-020-716-3  
Description: (1:5360) from US09020716.seq  
Perfect Score: 5360  
N.A. Sequence: 1 CTAATGTAAGCGTTAATA... ATTTCGCGAAAGTGGCAC 5360  
N.A. Comp: GATTTAACATTCGAATTAT... TAAAGGCGCTTTACCGGTG  
\*\*\*\*\*  
Scoring table:  
Gap K  
TABLE default  
Gap K

Nmatch STD : Dbase 0: Query 0  
Searched: 188442 seqs, 58026449 bases x 2  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: n-geneseq32  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39 40:part40  
Statistics: Mean 10.451: Variance 6.768: scale 1.546  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

Result No.	Score	Match	Query Length	ID	Description	Pred. No.
1	2222	41.5	5178	T49876	pTet-Splice	0.00e+00
2	2222	41.5	7474	T71320	Plasmid pTB73 encodin	0.00e+00
3	2220	41.4	6206	T49877	Autoregulatory vector	0.00e+00
4	2201	41.1	3681	Q13578	Plasmid pKSEL5.	0.00e+00
5	2202	41.1	4145	Q40279	Sequence of clone pS2	0.00e+00
6	2203	41.1	5314	Q13576	Plasmid pMTV1 contain	0.00e+00
7	2202	41.1	5356	T43794	Plasmid pPRIPAT (rat	0.00e+00
8	2202	41.1	5534	T43137	pMIGIT sequence incl	0.00e+00
9	2198	41.0	3699	V14340	Plasmid pBSGFP expres	0.00e+00
10	2198	41.0	4277	Q40280	Sequence of clone pS2	0.00e+00
11	2196	41.0	4539	Q87347	Plasmid pMTV1.	0.00e+00
12	2200	41.0	4792	Q84696	Plasmid GS contg. N.	0.00e+00
13	2200	41.0	5042	Q84694	Plasmid glucoamylase	0.00e+00

SUMMARIES







[illegible]

[illegible]







Db	4765	atcgtttgcagaagaatgcttgccgcagtggttatcactcatgcttattgacgacactgcgt	4824
Qy	4814	ATCGTTGTGAGAAGTAAGTTGSCGSCAGTGTATCAGTCATGCTTATGCGACGACTGCTAT	4873
Db	4825	aattctcttactgtcatgccatccgttaagatgctttctctgactagtgactcaacc	4884
Qy	4874	AA TTCCTTAC TGT CATG CATCCG TAAGATGCTTTCTGTGAC TGGT GAGTACTCAACC	4933
Db	4885	aagtcattctgaagaatagtgatgacgcgcagcagtgctcttaccgcgcgtcaatacgg	4944
Qy	4934	AAGTCATCTCTGAGAAATAGTGTATGGCGGACCGAGCTTGCTCTTGGCCGCGTCAATACGG	4993
Db	4945	gataataccgcgcacatagcaagaactttaaagtctcatcattggaagaacctcttcg	5004
Qy	4994	GATAA TACCGG GGCACATAGCAGACCTTTAAAGTGTCTCATTCATGTGAAAACGTTCTTGG	5053
Db	5005	ggcggaaaactctcagaggtatctaccgctgtgttaagatcccaattcgatgaacctctgt	5054
Qy	5054	GGCGGAAACCTCTCAAGGATCTTACCGCTGTGTGAGATCCAGTTCGATGTACCCACTCGT	5113
Db	5065	gcacccaactaatcttcagcatcttttactcttcacacagctttctgggtgaacaaacaa	5124
Qy	5114	GCACCCCAACTGATCTTCAGCATCTTTTACTTTCCACAGCGTTTCTGGTGAGCAAAAACA	5173
Db	5125	ggaagcgaataatgcgcgaataaagggaataaggcgcaacaggaataattgaataactata	5184
Qy	5174	GGAGGCAAAATGCCGCAAAAAAGGGAATAGGGGGACACGGAAATGTTGAATACTCATTA	5233
Db	5185	ctcttccttttcaattatttaagcattattcacgggttatgtctcatgagcggtatc	5244
Qy	5234	CTCTTCCCTTTTCAAATATTGGAAGCATTTATCAGGGTATTGTCTCATGAGCGGATAC	5293
Db	5245	atatttgaatgtatttgaataataataacuaatagagggttcgcgcgcacatttcccgcgaata	5304
Qy	5294	ATATTGAAATGATTAGAAAAATAACAAATAAGGSGTTCGCGGCACATTTCGCCGGAATA	5353
Db	5305	gtgccac	5311
Qy	5354	GTGCCAC	5360

```

RESULT 7
ID T43794 standard; DNA: 5356 BP.
AC T43794;
DT 18-FEB-1997 (first entry)
DE Plasmid pRIPHA1 (rat insulin promoter-human IAPP transgene).
KW Type II diabetes mellitus; transgenic animal model; pancreas;
KW islet; beta cell; islet amyloid polypeptide; IAPP; hyperglycemia;
KW glycosuria; diabetic glomerulosclerosis; plasmid pRIPHA1; ds;
KW cyclic.
KW Chimeric Homo sapiens;
OS Chimeric Rattus sp.;
OS Chimeric synthetic.
FH Key Location/Qualifiers
FT promoter 719..1594
FT CDS
FT /tag= a
FT /label= RIP-II_promoter
FT /note= "rat insulin II promoter"
FT 1589..1856
FT /tag= b
FT /label= IAPP
FT /note= "human IAPP coding sequence (Claim 7)"
FT intron 1861..2580
FT /tag= c
FT /label= Albumin_intron
FT /note= "human albumin intron 1"
FT 2575..3119
FT terminator
FT /tag= d
FT /label= Poly-A
FT /note= "human GAPDH gene poly-A and RNA
FT termination region"
PD WO9637612-A1.
PN 28-NOV-1996.

```





























CP 1200 00AAGGUGCAATTACGCTTACTTAAAGGAGACAAAGCTG: 4159

Search completed: Sun Mar 14 23:33:08 1999  
Job time: 1057 secs.



\*\*\*\*\*

**MAPSRELH**

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, BioComputing Research Unit  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sun Mar 14 20:54:50 1999; Maspar time 6810.13 Seconds  
1409.454 Million cell updates/sec

Tabular output not generated.

Title: >US-09-020-716-3  
Description: (1-5360) from US09020716 seq  
Perfect Score: 5360  
N.A. Sequence: 1 CTAAATGTAAGCGTTAATAATTCCCGGAAAGTCCAC 5360  
Comp: TAAAGGGCGTTTTCACGGTG

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: emb1-est56  
1:emb-est1 2:emb-gss1 3:emb-gss2 4:emb-gss3  
Database: genbank-est109  
5:gb-est1 6:gb-est10 7:gb-est11 8:gb-est12 9:gb-est13  
10:gb-est14 11:gb-est15 12:gb-est16 13:gb-est17  
14:gb-est18 15:gb-est19 16:gb-est2 17:gb-est20  
18:gb-est21 19:gb-est3 20:gb-est4 21:gb-est5 22:gb-est6  
23:gb-est7 24:gb-est8 25:gb-est9 26:gb-gss1 27:gb-gss2  
28:gb-gss3 29:gb-gss4

Statistics: Mean 12.787; Variance 2.653; scale 4.820

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	664	12.4	711	26	AG004608 Homo sapiens genomic D	0.00e+00
2	664	12.4	711	27	AG010947 Homo sapiens genomic D	0.00e+00
3	657	12.3	721	27	AG011001 Homo sapiens genomic D	0.00e+00
4	657	12.3	721	26	AG004662 Homo sapiens genomic D	0.00e+00
5	630	11.8	698	26	AG003787 Homo sapiens genomic D	0.00e+00
6	630	11.8	698	27	AG009976 Homo sapiens genomic D	0.00e+00
7	635	11.8	718	26	AG004363 Homo sapiens genomic D	0.00e+00
8	635	11.8	718	27	AG004489 Homo sapiens genomic D	0.00e+00
9	624	11.5	696	26	AG003785 Homo sapiens genomic D	0.00e+00
10	624	11.5	696	27	AG003786 Homo sapiens genomic D	0.00e+00
11	624	11.5	720	26	AG013858 Homo sapiens genomic D	0.00e+00
12	624	11.5	720	27	AG006062 Homo sapiens genomic D	0.00e+00
13	619	11.5	747	26	AG007052 Homo sapiens genomic D	0.00e+00

C	14	609	11.4	644	28	A0115013	CIT-HSP-2324C24.TF.CIT	0.00e+00
C	15	603	11.3	692	27	A0040352	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	16	603	11.3	695	26	AG002183	Homo sapiens genomic D	0.00e+00
C	17	598	11.2	748	17	A1124281	1.53 Drosophila 8-12 h	0.00e+00
C	18	593	11.1	654	27	A0074298	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	19	585	11.1	654	26	B78704	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	20	594	11.1	690	27	AG009464	Homo sapiens genomic D	0.00e+00
C	21	597	11.1	703	26	AG001761	Homo sapiens genomic D	0.00e+00
C	22	588	11.0	652	28	A108653	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	23	583	10.9	594	26	C17B6	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	24	583	10.9	666	27	A0079096	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	25	577	10.8	689	28	AG014394	Homo sapiens genomic D	0.00e+00
C	26	566	10.6	642	28	A0113213	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	27	566	10.6	682	27	AG074693	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	28	557	10.4	615	28	A012424	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	29	559	10.4	742	28	A0108392	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	30	551	10.3	604	26	AG002706	Homo sapiens genomic D	0.00e+00
C	31	550	10.3	634	27	A0038010	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	32	549	10.2	661	27	A0057603	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	33	545	10.2	690	27	A0029299	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	34	545	10.2	690	27	A0040787	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	35	544	10.1	637	27	A0077360	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	36	540	10.1	637	27	A0077360	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	37	534	10.0	594	26	B71605	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	38	536	10.0	636	28	A011342	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	39	531	9.9	549	27	A003191	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	40	530	9.9	630	27	A003238	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	41	526	9.8	743	26	AG002826	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	42	526	9.8	540	27	A0074307	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	43	526	9.8	594	26	AG007826	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	44	525	9.8	500	28	A011411	CIT-HSP-2324K21.TF.CIT	0.00e+00
C	45	527	9.8	629	27	AG011357	Homo sapiens genomic D	0.00e+00

ALIGNMENTS

RESULT 1 AG004608 711 bp DNA GSS 29-JAN-1998  
LOCUS Homo sapiens genomic DNA, 21q region, clone: S594BG38 genomic  
DEFINITION survey sequence.  
ACCESSION AG004608  
NID Q2822058  
KEYWORDS GSS.  
SOURCE Homo sapiens DNA, clone: S594BG38.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates, Catarrhini, Hominoidea, Hominidae, Homo  
REFERENCE 1 (bases 1 to 711)  
Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.  
AUTHORS Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.  
TITLE Published Only in Database (1998) in press  
JOURNAL  
REFERENCE 2 (bases 1 to 711)  
Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.  
AUTHORS Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.  
TITLE Direct Submission  
JOURNAL  
TITLE Submitted (29-JAN-1998) to the DDBJ/EMBL/GenBank databases.  
Masahira Hattori, Kitasato University, Department of Science, JSI  
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan  
(E-mail: hattori@jsgc.ims.u-tokyo.ac.jp, Tel: 0427-78-5732,  
Fax: 0427-78-9561)

FEATURES  
Location/Qualifiers  
1..711  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/clone="S594BG38"  
/map="21q"

BASE COUNT 169 a 188 c 172 g 182 t

ORIGIN

Query Match 12.4%; Score 664; DB 26; Length 711;  
Best Local Similarity 99.6%; Pred No. 0; Mismatches 0; Indels 3; Gaps 3;  
Matches 682; Conservative

```

28  TATATAATACGATACGACGAGGCTTACATCTGGGCTCAAGTGGGAAATACATACGCG 87
37  4454  TATATAATACGATACGACGAGGCTTACATCTGGGCTCAAGTGGGAAATACATACGCG 4512
38  AGAGCTACGCTACGCGCTGACATATATACGCAATATAAGGAGAGGAGGAGGAGGAG 147
43  4513  AGAGCTACGCTACGCGCTGACATATATACGCAATATAAGGAGAGGAGGAGGAGGAG 4572
44  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4632
45  4633  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4692
46  AGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 267
47  4693  AGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 4752
48  CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327
49  4753  CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4812
50  AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 447
51  4813  AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4872
52  TAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 507
53  4873  TAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4932
54  CAAGTAAATACGATACGATACGATACGATACGATACGATACGATACGATACGATACG 567
55  4933  CAAGTAAATACGATACGATACGATACGATACGATACGATACGATACGATACGATACG 4992
56  GSATAATACGATACGATACGATACGATACGATACGATACGATACGATACGATACG 627
57  4993  GSATAATACGATACGATACGATACGATACGATACGATACGATACGATACGATACG 5052
58  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 686
59  5053  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5112
60  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 711
61  5112  GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5176
62  AG010947  711 bp  DNA  28-JUL-1998  GSS
63  DEFINITION  Homo sapiens genomic DNA, 21q region, clone: 879612X91, genome
64  survey sequence.
65  ACCESSION  AG010947
66  KEYWORDS  GSS.
67  SOURCE  Homo sapiens
68  ORGANISM  Homo sapiens
69  Eukaryota; Chordata; Vertebrata; Mammalia; Eutheria;
70  Primates; Catarrhini; Hominidae; Homo.
71  1 (bases 1 to 711)
72  Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
73  Homo sapiens genomic DNA, chromosome 21q
74  Published only in Database (1998) In press
75  2 (bases 1 to 711)
76  Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
77  Direct Submission
78  Submitted (28-JUL-1998) to the DDBJ/EMBL/GenBank databases
79  Masahiro Hattori, Kitasato University Department of Science, 1st
80  Supporting Laboratory, Kitasato University, Sagami-cho 228, Japan
81  (E-mail:hattori@life.kit.ac.jp) Tel: 042-78-9922
82  Fax: 042-78-9961
83  Location/Qualifiers
84  FEATURES
85  CDS
86  1..711
87  /organism="Homo sapiens"
88  /db_xref="taxon:9606"
89  /chromosome="21"
90  /clone="S594HG38"
91  /map="21q"
92  159 a 188 c 172 g 192 t
93  ORIGIN
94  Query Match 12.4% Score 664.188 27.126 27.126
95  Best local similarity 99.68% Pred. No. 0.000000
96  Matches 682; Conservative 0; Mismatches 6; Indels 0; Gaps 0
97  Db 28 TATATAATACGATACGAGGAGGCTTACATCTGGGCTCAAGTGGGAAATACATACGCG 182
98  QY 4454 TATATAATACGATACGAGGAGGCTTACATCTGGGCTCAAGTGGGAAATACATACGCG 441
99  Db 88 AGAGCTACGCTACGCGCTGACATATATACGCAATATAAGGAGAGGAGGAGGAGGAG 143
100 QY 4513 AGAGCTACGCTACGCGCTGACATATATACGCAATATAAGGAGAGGAGGAGGAGGAG 443
101 Db 148 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 207
102 QY 4573 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 457
103 Db 208 AGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 267
104 QY 4633 AGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 463
105 Db 268 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327
106 QY 4693 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 472
107 Db 328 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 387
108 QY 4753 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 472
109 Db 488 CAAGTAAATACGATACGATACGATACGATACGATACGATACGATACGATACGATACG 447
110 QY 4933 CAAGTAAATACGATACGATACGATACGATACGATACGATACGATACGATACGATACG 492
111 Db 568 GSATAATACGATACGATACGATACGATACGATACGATACGATACGATACGATACG 627
112 QY 4993 GSATAATACGATACGATACGATACGATACGATACGATACGATACGATACGATACG 5052
113 Db 628 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 686
114 QY 5053 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5112
115 Db 687 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 711
116 QY 5112 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5176
117 3 AG010947 721 bp  DNA  GSS
118 DEFINITION  Homo sapiens genomic DNA, 21q region, clone: 879612X91, genome
119 survey sequence.
120 ACCESSION  AG010947
121 KEYWORDS  GSS.
122 SOURCE  Homo sapiens
123 ORGANISM  Homo sapiens
124 Eukaryota; Chordata; Vertebrata; Mammalia; Eutheria;
125 Primates; Catarrhini; Hominidae; Homo.

```

Wed Mar 17 09:43:39 1999

1 (bases 1 to 721)  
 Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.  
 Homo sapiens genomic DNA, chromosome 21q  
 Published only in Database (1998) in press  
 2 (bases 1 to 721)  
 Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.  
 Direct Submission  
 Submitted (29-JUL-1998) to the DDBJ/EMBL/GenBank databases.  
 Masahira Hattori, Kitasato University, Department of Science, JST  
 Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan  
 (E-mail:hattori@scs.kit.ac.jp, Tel:0427-78-9732,  
 Fax:0427-78-9561)  
 Location/Qualifiers  
 1..721  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="21"  
 /clone="879G12X91"  
 /map="21q"  
 /map="21q"  
 BASE COUNT 195 a 163 c 175 g 179 t 9 others  
 ORIGIN  
 Query Match 12.3%, Score 657, DB 27, Length 721:  
 Best Local Similarity 98.7%; Pred. No. 0.00e-00;  
 Matches 683; Conservative 0; Mismatches 5; Indels 4; Gaps 4;  
 Db 32 CATCTTACGGATGGCATGACAGTAAGAGAAATTTATGAGTGGTGGCATAACCATGATGAT 91  
 Cp 4905 CATCTTACGGATGGCATGACAGTAAGAGAAATTTATGAGTGGTGGCATAACCATGAT 4846  
 Db 92 AACACTGGGGCCAACTTACTTCTGACAGCACTGCGAGGACCGAAGGAGCTTAACCGCTTTT 151  
 Cp 4845 AACACTGGGGCCAACTTACTTCTGACAGCACTGCGAGGACCGAAGGAGCTTAACCGCTTTT 4786  
 Db 152 TTGCACAACTGGGGGATCATGTAACCTGCGCTTGCATGCTGATGCTGATGCTGATGAA 211  
 Cp 4785 TTGCACAACTGGGGGATCATGTAACCTGCGCTTGCATGCTGATGCTGATGAA 4726  
 Db 212 GGCATACCAAAAG 271  
 Cp 4725 GGCATACCAAAAG 4666  
 Db 272 AAACATTAACCTGGGAACTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 331  
 Cp 4665 AAACATTAACCTGGGAACTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 4796  
 Db 332 GAGGCGGATTAAGATTTGACAGGACCACTTCTGAGGTCGGGCGCTTGGCTGGCTGGCTT 391  
 Cp 4605 GAGGCGGATTAAGATTTGACAGGACCACTTCTGAGGTCGGGCGCTTGGCTGGCTGGCTT 4546  
 Db 392 GCTGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 451  
 Cp 4545 GCTGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 4486  
 Db 452 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 511  
 Cp 4485 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 4426  
 Db 512 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 570  
 Cp 4425 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 4366  
 Db 571 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 629  
 Cp 4365 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 4306  
 Db 629 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 687  
 Cp 4305 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 4246  
 Db 687 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 719  
 Cp 4245 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 4214

4  
 AG004662 721 bp DNA GSS 30-JAN-1998  
 Homo sapiens genomic DNA, 21; region, clone: g79g12x91, genomic  
 survey sequence.  
 AG004662  
 GSS  
 52826191  
 Homo sapiens DNA, clone:879G12X91.  
 SOURCE  
 Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.  
 Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
 Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 721)  
 Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.  
 Homo sapiens genomic DNA, chromosome 21q  
 Published only in Database (1998) in press  
 2 (bases 1 to 721)  
 Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.  
 Direct Submission  
 Submitted (30-JAN-1998) to the DDBJ/EMBL/GenBank databases.  
 Masahira Hattori, Kitasato University, Department of Science, JST  
 Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan  
 (E-mail:hattori@scs.kit.ac.jp, Tel:0427-78-9732,  
 Fax:0427-78-9561)  
 Location/Qualifiers  
 1..721  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="21"  
 /clone="879G12X91"  
 /map="21q"  
 /map="21q"  
 BASE COUNT 195 a 163 c 175 g 179 t 9 others  
 ORIGIN  
 Query Match 12.3%, Score 657, DB 26, Length 721:  
 Best Local Similarity 98.7%; Pred. No. 0.00e-00;  
 Matches 683; Conservative 0; Mismatches 5; Indels 4; Gaps 4;  
 Db 32 CATCTTACGGATGGCATGACAGTAAGAGAAATTTATGAGTGGTGGCATAACCATGATGAT 91  
 Cp 4905 CATCTTACGGATGGCATGACAGTAAGAGAAATTTATGAGTGGTGGCATAACCATGATGAT 4846  
 Db 92 AACACTGGGGCCAACTTACTTCTGACAGCACTGCGAGGACCGAAGGAGCTTAACCGCTTTT 151  
 Cp 4845 AACACTGGGGCCAACTTACTTCTGACAGCACTGCGAGGACCGAAGGAGCTTAACCGCTTTT 4796  
 Db 152 TTGCACAACTGGGGGATCATGTAACCTGCGCTTGCATGCTGATGCTGATGCTGATGAA 211  
 Cp 4785 TTGCACAACTGGGGGATCATGTAACCTGCGCTTGCATGCTGATGCTGATGAA 4726  
 Db 212 GGCATACCAAAAG 271  
 Cp 4725 GGCATACCAAAAG 4666  
 Db 272 AAACATTAACCTGGGAACTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 331  
 Cp 4665 AAACATTAACCTGGGAACTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 4796  
 Db 332 GAGGCGGATTAAGATTTGACAGGACCACTTCTGAGGTCGGGCGCTTGGCTGGCTGGCTT 391  
 Cp 4605 GAGGCGGATTAAGATTTGACAGGACCACTTCTGAGGTCGGGCGCTTGGCTGGCTGGCTT 4546  
 Db 392 GCTGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 451  
 Cp 4545 GCTGATAAAATCTGGAGGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 4486  
 Db 452 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 511  
 Cp 4485 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 4426  
 Db 512 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 570  
 Cp 4425 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 4366  
 Db 571 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 629  
 Cp 4365 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 4306  
 Db 629 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 687  
 Cp 4305 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 4246  
 Db 687 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 719  
 Cp 4245 GATGTTAAGCCCTCCCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 4214













Venter, J.C.  
Use of a random human BAC End Sequence Database for Sequence-Ready  
Map Building  
Unpublished (1998)

Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
7712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdamas@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
http://www.tigr.org/tcdb/human/bac\_end\_search/bac\_end\_search.html.  
Seq primer: M13 Reverse  
Class: BAC ends.

FEATURES  
Source  
1. .644

/organism="Homo sapiens"  
/note="Vector: pBeloBAC11; Site\_1: HindIII; Site\_2:  
HindIII"  
/db\_xref="taxon:9606"  
/clone="2374C24"  
/clone\_lib="Cif-Hsp"  
/sex="Male"  
/cell\_type="Sperm"  
135 a 183 c 145 g 179 t 2 others

BASE COUNT 135 a 183 c 145 g 179 t 2 others  
ORIGIN

Query Match 11.4%; Score 609; DB 28; Length 644;  
Best Local Similarity 99.8%; Pred. No. 0.00e+00;  
Matches 615; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

5 TCACTGCGCTGCTTTTACAGCTGCTGAGTGGGAAACDCTGGCGTTACCCAACTTAAT 64  
Cp 618 TCATGCGCGCTGCTTTTACAGCTGCTGAGTGGGAAACDCTGGCGTTACCCAACTTAAT 559  
Db 65 CGCCTTCAGCAGCATCCGCTTTTCGCCAGCTGGCGTAATAGCGAAGCGCGGACGAT 124  
Cp 558 CGCCTTCAGCAGCATCCGCTTTTCGCCAGCTGGCGTAATAGCGAAGCGCGGACGAT 499  
Db 125 CGCCTTCAGCAGCATCCGCTTTTCGCCAGCTGGCGTAATAGCGAAGCGCGGACGAT 193  
Cp 498 CGCCTTCAGCAGCATCCGCTTTTCGCCAGCTGGCGTAATAGCGAAGCGCGGACGAT 439  
Db 184 TTAAGCGGGGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 243  
Cp 438 TTAAGCGGGGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 379  
Db 244 GCGGCGGCTGCTTCGCTTTTCGCTTTTCGCTTTTCGCTTTTCGCTTTTCGCTTTTCGCT 308  
Cp 378 GCGGCGGCTGCTTTTCGCTTTTCGCTTTTCGCTTTTCGCTTTTCGCTTTTCGCTTTTCGCT 319  
Db 304 CAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCGGATTAGTGGTTCGATAGCGACGAC 463  
Cp 318 CAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCGGATTAGTGGTTCGATAGCGACGAC 259  
Db 364 CCCAAAAAATCTGATTAGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 423  
Cp 259 CCCAAAAAATCTGATTAGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 199  
Db 424 TTTCGCCCTTTACGCTTGGAGTCCACGCTTCTTTTAATAGTGGACCTTTTCCAACTGGA 483  
Cp 198 TTTCGCCCTTTACGCTTGGAGTCCACGCTTCTTTTAATAGTGGACCTTTTCCAACTGGA 139  
Db 484 ACAAGCTCAACCTATCTCGGCTTATCTTTTGGATTATAGGGGATTTTGGCGATTTCG 543  
Cp 138 ACAAGCTCAACCTATCTCGGCTTATCTTTTGGATTATAGGGGATTTTGGCGATTTCG 79  
Db 544 GCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGGAATTTTAAACAAAAA 603  
Cp 78 GCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGGAATTTTAAACAAAAA 19

TITLE  
JOURNAL  
COMMENT

BASE COUNT 194 a 171 c 176 g 194 t 12 others  
ORIGIN

Query Match 11.5%; Score 619; DB 26; Length 747;  
Best Local Similarity 96.3%; Pred. No. 0.00e+00;  
Matches 683; Conservative 0; Mismatches 17; Indels 9; Gaps 9;

2b 41 ATCTGGGCTGCTGCTGAAGCAGTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGC 100  
QY 4057 ATCTGGGCTGCTGCTGAAGCAGTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGC 4126  
Db 101 AAACAAAACATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 159  
QY 412 AAACAAAACATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 4186  
Db 160 AAAAAGGATCTCAAGAGATCTTTTGGATCTTTTACGGGGTCTGACGCTCAGTGGAAAC 219  
QY 4187 AAAAAGGATCTCAAGAGATCTTTTGGATCTTTTACGGGGTCTGACGCTCAGTGGAAAC 4246  
Db 220 GAAAGCTCAGCTTAAGGATTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 279  
QY 4247 GAAAGCTCAGCTTAAGGATTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 4306  
Db 280 CTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAAGTTGGTCT 339  
QY 4307 CTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATAGTAAAGTTGGTCT 4366  
Db 340 GAAAGTACCAATGCTTAATAGTGGGAGCTATCTCAGGGATCTCTCTATTTCGTTCA 399  
QY 4367 GAAAGTACCAATGCTTAATAGTGGGAGCTATCTCAGGGATCTCTCTATTTCGTTCA 4426  
Db 400 TCATAGTTCGCTGACTCCCGCTGCTGGTAGATACTACATACGAGGGGCT-ACCATCT 458  
QY 4427 TCATAGTTCGCTGACTCCCGCTGCTGGTAGATACTACATACGAGGGGCT-ACCATCT 4486  
Db 459 GAGGCTAGTCTCAATGATACGGGAGACCCACGCTCAGCGGCTCCAGATTATACGA 518  
QY 4487 GAGGCTAGTCTCAATGATACGGGAGACCCACGCTCAGCGGCTCCAGATTATACGA 4546  
Db 519 AATAAATAGCTAGCGGAGAGGGGCTCAGAGGCA-AGTGGTCTCTGTAATTTATCCGGCTC 577  
QY 4547 AATAAATAGCTAGCGGAGAGGGGCTCAGAGGCA-AGTGGTCTCTGTAATTTATCCGGCTC 4605  
Db 578 CATCCARTCTATTAATTTGTCGGGAGAGCTAGAGTAAGTAACTTTCGCGAGTTAATAGT 637  
QY 4606 CATCCARTCTATTAATTTGTCGGGAGAGCTAGAGTAAGTAACTTTCGCGAGTTAATAGT 4663  
Db 638 TTGGGCAAGCTGTTGNCATTTGTTGGAGGATCCTGCTGACGCGCTCCCTTTGGTAA 697  
QY 4664 TTGGGCAAGCTGTTGNCATTTGTTGGAGGATCCTGCTGACGCGCTCCCTTTGGTAA-T 4722  
Db 698 GATTCATCTTTCAGTTCGGGCTTCCCAACGATCAAGGGGGTTACATGAT 746  
QY 4723 GATTCATCTTTCAGTTCGGGCTTCCCAACGATCAAGGGGGTTACATGAT 4769

RESULT 14 A0115013 644 bp DNA GSS 28-AUG-1998  
LOCUS CII-HSP-2374C24.TP CII-HSP Homo sapiens genomic clone 2374C24,  
DEFINITION genomic survey sequence.

ACCESSION A0115013  
NID 33491134  
KEYWORDS GSS  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 644)  
AUTHORS Adams, M.L., Founley, S.D., Zhao, S., Linher, K., Golden, K.,  
Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, M. and



\*\*\*\*\*  
WAPR 17 1999  
\*\*\*\*\*  
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit,  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nn n.a - n.a database search, using Smith-Waterman algorithm  
Run on: Mon Mar 15 01:29:29 1999. Maspar time 8427.16 Seconds  
1545.419 Million cell updates/sec

Tabular output not generated

Title: >US-09-020-716-4  
Description: (1:5511) from USG020716 seq  
Perfect Score: 5511  
N.A. Sequence: 1 TCGCGCGTTTGGTGATGAC . . . TATCAGGAGGCCCTTTGTC 5511  
Comp: ACGCGCAAGCAGCTATG . . . ATAGTGTCCGGCAAGCAG

Scoring table:  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 602357 seqs, 1181590623 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: emb157  
1:em\_ba 2:em\_fun 3:em\_htrg 4:em\_hum1 5:em\_hum2 6:em\_in  
7:em\_om 8:em\_ov 9:em\_ov 10:em\_pat 11:em\_ph 12:em\_pl  
13:em\_ro 14:em\_un 15:em\_vl

Database: genbank110  
16:gb\_bal 17:gb\_ba2 18:gb\_htrg 19:gb\_in 20:gb\_om 21:gb\_ov  
22:gb\_pat 23:gb\_ph 24:gb\_pl 25:gb\_pl2 26:gb\_pr1  
27:gb\_pr2 28:gb\_pr3 29:gb\_ro 30:gb\_st 31:gb\_sts 32:gb\_sy  
33:gb\_un 34:gb\_vl

Statistics: Mean 12.522; Variance 6.223; scale 2.012

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARY				SUMMARY			
Result No.	Score	Match Length	ID	Description	Pred. No.		
1	2243	40.7	2689	32 Cloning vector pUC-JK	0.00e+00		
C	2	2241	40.7	4102 32 Cloning vector pBRV-2	0.00e+00		
	3	2242	40.7	4506 32 Cloning vector pCAT-PR	0.00e+00		
	4	2243	40.7	4884 32 Sequence 1 from patent	0.00e+00		
	5	2243	40.7	4883 32 Sequence 1 from patent	0.00e+00		
C	6	2242	40.7	4883 32 DNA for transforming m	0.00e+00		
	7	2242	40.7	4883 32 Sequence 4 from patent	0.00e+00		
	8	2242	40.7	4883 32 Cloning vector pBRV-1	0.00e+00		
	9	2241	40.7	5399 22 Cloning vector pBRV-1	0.00e+00		
C	10	2241	40.7	5399 22 DNA for transforming m	0.00e+00		
	11	2241	40.7	5399 22 Sequence 1 from patent	0.00e+00		
	12	2242	40.7	5620 32 Sequence 4 from patent	0.00e+00		
	13	2242	40.7	5620 32 pVE108 plasmid DNA	0.00e+00		

RESULT	LOCUS	1	AF016541	2689 bp	DNA	circular	SYN	04-SEP-1997
DEFINITION	Cloning vector pUC-JK, complete sequence.							
ACCESSION	AF016541							
NID	92353791							
KEYWORDS	Cloning vector pUC-JK.							
SOURCE	ORGANISM							
	Artificial sequence; cloning vectors.							
REFERENCE	1 (bases 1 to 2689)							
AUTHORS	Kim, J.J., Brynan, B.E. and Quackenbush, J.							
TITLE	pUC-JK cloning vector							
JOURNAL	Unpublished							
REFERENCE	2 (bases 1 to 2689)							
AUTHORS	Kim, J.J., Brynan, B.E. and Quackenbush, J.							
TITLE	Direct Submission							
JOURNAL	Submitted (30-JUL-1997) The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20878, USA							
FEATURES	Location/Qualifiers							
Source	1..2689							
	/organism="Cloning vector pUC-JK"							
	/db_xref="taxon:65020"							
	/plasmid="pUC-JK"							
	/note="high copy number plasmid derived from pUC19"							
	/complement(146..472)							
	/codon_start=1							
	/transl_table=11							
	/product="beta-galactosidase alpha peptide"							
	/db_xref="PIR:1233792"							
	/translation="MTITPFLVNSTCTPFGTPTGNSLAVVLAFFRWENFGVTQL NLAHAFFAFWAPSEAFPTPPSQRLSNGEWFLMRVLLFLGGSIRIWCITLST ICSDA"							
	/complement(395..455)							
	/note="multiple cloning site adapted from pUC19 for BstXI cloning; contains restriction sites for EcoRI, BstXI, KpnI, SmaI, BamHI, PstI, SalI, BstXI and HindIII"							
	/cdo="Region: multiple cloning site"							

ALIGNMENTS

14	2342	40.7	5620	22	AR007513	Sequence 3 from patent	0.00e+00
15	2342	40.7	5642	22	170384	Sequence 2 from patent	0.00e+00
16	2342	40.7	5647	16	BACNNR	Bacillus stearothermophilus	0.00e+00
17	2342	40.7	6291	22	A60111	Sequence 4 from patent	0.00e+00
18	2342	40.7	6376	22	133114	Sequence 5 from patent	0.00e+00
19	2342	40.7	6376	22	A23633	Plasmid pVE149	0.00e+00
20	2342	40.7	6555	22	A8607512	Sequence 2 from patent	0.00e+00
21	2342	40.7	7049	22	A23998	pB4 expression vector	0.00e+00
22	2342	40.7	11490	32	CVU580929	Cloning vector pBACE4	0.00e+00
23	2342	40.7	16297	32	CVU75991	Cloning vector pBACE4	0.00e+00
24	2342	40.7	19500	32	CVU75992	Cloning vector pBACE4	0.00e+00
25	2342	40.7	2860	32	AF017063	Cloning vector pMECA	0.00e+00
26	2339	40.6	3164	23	BPLVSPR	Bacteriophage phi-X174	0.00e+00
27	2340	40.6	3409	32	CVU94951	Cloning vector pZiP1	0.00e+00
28	2338	40.6	3563	32	XU04610	Cloning vector pUCM 9	0.00e+00
29	2340	40.6	4392	32	CVU47234	Cloning vector pAdvant	0.00e+00
30	2338	40.6	4446	32	XU034310	Shuttle vector pCX111	0.00e+00
31	2340	40.6	4630	32	CVU94950	Cloning vector pBAP2/S	0.00e+00
32	2338	40.6	4630	32	CVU35316	Cloning vector pBAP2	0.00e+00
33	2338	40.6	5008	32	XU017501	Cloning vector pX1018G	0.00e+00
34	2338	40.6	5154	32	XU017502	Cloning vector pX1018G	0.00e+00
35	2338	40.6	5257	32	XU034320	Cloning vector pUC-GM	0.00e+00
36	2338	40.6	5645	32	AF025362	Cloning vector pUC-GM	0.00e+00
37	2338	40.6	5862	32	XU034311	E.coli/Kluyveromyces 1	0.00e+00
38	2338	40.6	6461	32	XU034315	E.coli/Kluyveromyces 1	0.00e+00
39	2340	40.6	6687	32	AB012282	Shuttle vector pSVbeta	0.00e+00
40	2338	40.6	6894	32	XU024335	Cloning vector pSVbeta	0.00e+00
41	2338	40.6	7561	32	XU034313	E.coli/Kluyveromyces 1	0.00e+00
42	2339	40.6	9208	32	AF064057	Expression vector pYZ8	0.00e+00
43	2339	40.6	9211	32	AF064056	Expression vector pYZ4	0.00e+00
44	2339	40.6	9215	32	AF049083	Expression vector pYZ1	0.00e+00
45	2339	40.6	9956	32	AF049084	Expression vector pYZ3	0.00e+00



```

DRYDIALATRDIAEELGGSWADRFVLYGIAAPDSORIAFYRLDEFF"
3485..3490
/Note="hsv thymidine kinase polyA signal"
polyA_signal
3498..3503
/Note="hsv thymidine kinase polyA signal"
polyA_signal
3572..3605
/Note="loxP element"
misc_feature
3674..3802
/Note="multiple cloning site (ClaI, BglII, Pst-EII, NotI)"
misc_feature

BASE COUNT 967 a 1059 c 1085 g 991 t
ORIGIN

Query Match 40.7%; Score 2341; E 37; Length 4102;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 GAGGAAGGCGCTCGTATACGCCATATTTTATAGGTTAAATGTCATGATATATATGTTT 60
|||||
Cp 5511 GAGGAAGGCGCTCGTATACGCCATATTTTATAGGTTAAATGTCATGATATATGTTT 5452
|||||
Db 61 CTTAGAGCTGAGTGGGCGCTTTTCGCGGAAATGTAAGGAAAGGCGCTATTTGTTTATTTT 120
|||||
Cp 5451 CTTAGAGCTGAGTGGGCGCTTTTCGCGGAAATGTAAGGAAAGGCGCTATTTGTTTATTT 5362
|||||
Db 121 TCTAAATACATTCAAATATGATTCGCTTCATGAGACAATAACCTGATAAATGCTTCAAT 180
|||||
Cp 5391 TCTAAATACATTCAAATATGATTCGCTTCATGAGACAATAACCTGATAAATGCTTCAAT 5332
|||||
Db 181 AATATGAAAGGAGAGATGATGATTAATCAACATTTGCGTGTGCGGCTTATTCGCTTTT 240
|||||
Cp 5331 AATATGAAAGGAGAGATGATGATTAATCAACATTTGCGTGTGCGGCTTATTCGCTTTT 5272
|||||
Db 241 TTGGGCGCATTTTGGCTTCCTTGTCTTTCCTCAGCCAGAAAGCGTGTGTAAGTAAGATG 300
|||||
Cp 5271 TTGGGCGCATTTTGGCTTCCTTGTCTTTCCTCAGCCAGAAAGCGTGTGTAAGTAAGATG 5212
|||||
Db 301 CTGAAGATCACTTGGGTGCGAGAGTGGTTTACATCACTGATCTCAACAGCGGTAGA 360
|||||
Cp 5211 CTGAAGATCACTTGGGTGCGAGAGTGGTTTACATCACTGATCTCAACAGCGGTAGA 5152
|||||
Db 361 TCGTTGAGAGTTTTCGCGCGGAGAAAGGTTTTCATATGATGATGATGATGATGATG 420
|||||
Cp 5151 TCGTTGAGAGTTTTCGCGCGGAGAAAGGTTTTCATATGATGATGATGATGATGATG 5092
|||||
Db 421 TATGTGGCGCGGTATTTATCCGCTATTTGACGCGCGGAGAAAGCACTGCGTCCGCGATAC 480
|||||
Cp 5091 TATGTGGCGCGGTATTTATCCGCTATTTGACGCGCGGAGAAAGCACTGCGTCCGCGATAC 5032
|||||
Db 481 ACTATCTCAGAAATCACTTGGTTGATGACTCACCAGTCCAGAAAAGCATCTTTACAGATG 540
|||||
Cp 5031 ACTATCTCAGAAATCACTTGGTTGATGACTCACCAGTCCAGAAAAGCATCTTTACAGATG 4972
|||||
Db 541 GCATGACAGTAAGAGAAATTTGCGAGTGGTGCATACCATGATGATGATGATGATGATG 600
|||||
Cp 4971 GCATGACAGTAAGAGAAATTTGCGAGTGGTGCATACCATGATGATGATGATGATG 4912
|||||
Db 601 ACTATCTCAGAAATCACTTGGTGGAGAGCGGAGAGCTTAACCGCTTTTTCGACAAATGG 660
|||||
Cp 4911 ACTATCTCAGAAATCACTTGGTGGAGAGCGGAGAGCTTAACCGCTTTTTCGACAAATGG 4852
|||||
Db 661 GGGATCATGTAATCTGCGCTTATGCTTGGGAAACCGGAGCTTAATGAGGATACCAACG 720
|||||
Cp 4851 GGGATCATGTAATCTGCGCTTATGCTTGGGAAACCGGAGCTTAATGAGGATACCAACG 4792
|||||
Db 721 ACGAGCGTACACCGAGTATGCTGTAGCAATGGCAACCAACGTTGCGCAAACTATTAAGTG 780
|||||
Cp 4791 ACGAGCGTACACCGAGTATGCTGTAGCAATGGCAACCAACGTTGCGCAAACTATTAAGTG 4732
|||||
Db 781 GCGAACTACTTACTTACTGCTTCCCGGCAACAAATTAATATGATGATGATGATGATGATG 840
|||||
Cp 4731 GCGAACTACTTACTTACTGCTTCCCGGCAACAAATTAATATGATGATGATGATGATG 4672
|||||
Db 841 TTGAGGAGCAACCTTCTGCGCTTCCCGGCTTCCCGGCTTCCCGGCTTCCCGGCTTCCCGGCT 900
|||||

```



[illegible]









```

Db 3542 ACATTCGAAAGAGATREFAAGTCTTCATGCTGGGAAA'AAAGACACAGCTGGTAGGGGT 3601
QY 4130 ACATTCGAAAGAGATREFAAGTCTTCATGCTGGGAAA'AAAGACACAGCTGGTAGGGGT 4229
Db 3662 ATACTCTTCCTCTTTTCAATATATGAGGCAATTAAGAGGATTAATGCTTATATATATATAT 4741
QY 5310 ATAGCTCTTCCTTTTCAATATATGAGGCAATTAAGAGGATTAATGCTTATATATATATAT 5409
Db 4742 TACATATTTGATGATTAATGAGGCAATTAAGAGGATTAATGCTTATATATATATATATAT 4841
QY 5370 TACATATTTGATGATTAATGAGGCAATTAAGAGGATTAATGCTTATATATATATATATAT 5469
Db 4802 AAAGTCCACCTGAGCTTACAGGATTAATGAGGCAATTAATGCTTATATATATATATATAT 4901
QY 5430 AAAGTCCACCTGAGCTTACAGGATTAATGAGGCAATTAATGCTTATATATATATATATAT 5469
Db 4862 GGTATCAGGAGGCTTGTGTC 4884
QY 5466 GGTATCAGGAGGCTTGTGTC 4911

RESULT 6
LOCUS A18053 4884 bp DNA PAT 05 APR 1994
DEFINITION DNA for transformed neoplasty-derived plants.
ACCESSION A18053
NID q513158
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 4883)
AUTHORS
JOURNAL
FEATURES
Source
misc_feature
promoter
CDS
polyA_site
misc_feature
BASE COUNT 1301 a 1256 c 1201 g 1125 t
ORIGIN
Query Match 40.7% Score 23423.39 E22 Length 4883
Best Local Similarity 100.0% Ident. No. 0.000000
Matches 22423 Conservative 0.000000 0.000000 0.000000 0.000000

Db 2642 GAAGAGCTGGGGTAATCAATGAGGCAATTAAGAGGATTAATGCTTATATATATATATATAT 2701
QY 3270 GAAGAGCTGGGGTAATCAATGAGGCAATTAAGAGGATTAATGCTTATATATATATATATAT 3369
Db 2702 ATTCACACACATACAGGAGGAGGATTAATGAGGCAATTAAGAGGATTAATGCTTATATATAT 2761
QY 3330 ATTCACACACATACAGGAGGAGGATTAATGAGGCAATTAAGAGGATTAATGCTTATATATAT 3389
Db 2762 AGCTAACTCACATTAATGCTTGTGTCATGCTTGTGTCATGCTTGTGTCATGCTTGTGTCAT 2821
QY 3390 AGCTAACTCACATTAATGCTTGTGTCATGCTTGTGTCATGCTTGTGTCATGCTTGTGTCAT 3449

```

















3271 CAAGCTTGGCGTAATCATGTGTCATAGCTGTTTCTGTGTGTAATTTGTTATTCGGCTCACAA 3330  
 3219 TTCCACAAACATACGAGCGGAAAGCATATAAGTGTAAAGCGTGGGTTGGCTTAATGAGTGA 3278  
 3331 TTCCACAAACATAGAGAGCGGAAAGCATATAAGTGTAAAGCGTGGGTTGGCTTAATGAGTGA 3390  
 3279 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3338  
 3391 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3450  
 3339 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3398  
 3451 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3510  
 3399 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3458  
 3511 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3570  
 3459 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3518  
 3571 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3630  
 3519 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3578  
 3631 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3690  
 3579 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3638  
 3691 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3750  
 3639 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3698  
 3751 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3810  
 3699 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3758  
 3811 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3870  
 3759 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3818  
 3871 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3930  
 3819 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3878  
 3931 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3990  
 3879 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3938  
 3991 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4050  
 3939 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3998  
 4051 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4110  
 3999 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4058  
 4111 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4170  
 4059 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4118  
 4171 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4230  
 4119 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4178  
 4231 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4290  
 4179 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4238  
 4291 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4350  
 4239 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4298  
 4351 GGTAACTACATTAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4410

4831 CAAGCGAGTACATGATCCCGCATGTTGTCACAAAGCGGTTAGCTCTTCGGTCTTC 4890  
 4779 CGATGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 4838  
 4891 CGATGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 4950  
 4839 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 4898  
 4951 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 5010  
 4899 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 4958  
 5011 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 5070  
 4959 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 5018  
 5071 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 5130  
 5019 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 5078  
 5131 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 5190  
 5079 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 5138  
 5191 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 5250  
 5139 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 5198  
 5251 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 5310  
 5199 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 5258  
 5311 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 5370  
 5259 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 5318  
 5371 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 5430  
 5319 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 5378  
 5431 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 5490  
 5379 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 5399  
 5491 ATAACTGCTGTCAGAGTAAAGTGGCGGAGTGTATCACTCATGTTATGGCAGCACTGC 5511

RESULT 11 182373 5399 bp DNA PAT 20-MAR-1998  
 LOCUS Sequence 1 from patent US 5712135.  
 DEFINITION 182373  
 ACCESSION g3210670  
 NID  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE  
 1. (bases 1 to 5399)  
 D'Halluin, K. and Gobel, E.  
 Process for transforming monocotyledonous plants  
 Patent: US 5712135-A 1 27-JAN-1998;  
 Location/Qualifiers  
 1..5399  
 /organism="unknown"  
 BASE COUNT 1359 a 1378 c 1369 g 1293 t  
 ORIGIN  
 Query Match 40.7%, Score 2241, DB 22 Length 5399;  
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
 Matches 2241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 3159 CAAGCTTGGCGTAATCATGTGTCATAGCTGTTTCTGTGTGTAATTTGTTATTCGGCTCACAA 3218



```

QY 3930 ACCAAGTCATTCTGAGAAATAGTGTATGCGGGGACCGAGTTGCTCTTGCGCGGGCTCAATA 5069
DB 5179 CCGGATAATATGCGGCGCCACATAGCAGAACTTTAAAGATGTTTAAATGATGAGAAAGCTTCT 5238
QY 5070 CCGGATAATATGCGGCGCCACATAGCAGAACTTTAAAGATGTTTAAATGATGAGAAAGCTTCT 5129
DB 5239 TGGGGCGGAAAAGCTGTCAGGAGATCTTACGAGTCTTACGAGTCTTACGAGTCTTACGAGTCT 5298
QY 5130 TGGGGCGGAAAAGCTGTCAGGAGATCTTACGAGTCTTACGAGTCTTACGAGTCTTACGAGTCT 5199
DB 5299 CTTGCAACCAAGTATCTTACGAGTCTTACGAGTCTTACGAGTCTTACGAGTCTTACGAGTCT 5358
QY 5190 CTTGCAACCAAGTATCTTACGAGTCTTACGAGTCTTACGAGTCTTACGAGTCTTACGAGTCT 5249
DB 5359 ACAGGAAGCAAAATGCGGCAAAAGGCAATAGGCGGCAATAGGCGGCAATAGGCGGCAATAGGCGG 5418
QY 5250 ACAGGAAGCAAAATGCGGCAAAAGGCAATAGGCGGCAATAGGCGGCAATAGGCGGCAATAGGCGG 5309
DB 5419 ATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5478
QY 5310 ATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5369
DB 5479 TACATTTTGAATGTTTATAGAAAATAGGCGGCAATAGGCGGCAATAGGCGGCAATAGGCGGCA 5538
QY 5370 TACATTTTGAATGTTTATAGAAAATAGGCGGCAATAGGCGGCAATAGGCGGCAATAGGCGGCA 5429
DB 5539 AAGTGGCAGCTGACGCTGTAAGAAAAGCAATATATCATGAGATTAATATATATATATATATAT 5598
QY 5430 AAGTGGCAGCTGACGCTGTAAGAAAAGCAATATATCATGAGATTAATATATATATATATATAT 5489
DB 5599 CGTATGCGAGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5620
QY 5490 CGTATGCGAGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5511

RESULT 13 A2334 5620 bp LNA PAT 01-APR-1995
LOCUS PVE108 plasmid DNA.
DEFINITION A2334
ACCESSION A2334
KEYWORDS 9904321
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
REFERENCE 1 (bases 1 to 5620)
AUTHORS STAMEN-SPECIFIC PROMOTERS FROM RICE
TITLE Patent. WO 9213956-A 9 20-AUG-1992.
JOURNAL Location/Qualifiers
FEATURES
    source
        1..5620
        /organism="Escherichia coli"
        /plasmid="Plasmid pVE108"
        /db_xref="taxon:562"
    misc_feature
        1..395
        /note="pUC18 derived sequence"
    misc_feature
        396..802
        /note="Agrobacterium T-DNA derived sequence: contains
        polyA site from nopaline gene"
    promoter
        1139..11683
        /note="tapetum-specific; Nicotiana tabacum derived
        sequence"
    promoter
        1684..2516
        /note="35S3; Callitower mosaic virus isolate Cab88-31
        derived sequence"
    gene
        2517..3068
        /gene="bar"
    CDS
        2517..3068
        /gene="bar"
        /codon_start=1
        /transl_table=1
        /product="phosphotriester transferase"
        /db_xref="F02:9904321"

```













\*\*\*\*\*  
[REDACTED]  
\*\*\*\*\*  
(TM)

Release 3.1A John F. Collins, Bioinformatics Research Unit,  
Copyright (c) 1993-1998 University of Edinburgh, UK  
Distribution rights by Oxford Molecular Ltd

MSearch\_nnn n.d. - n.d. database search, using Smith-Waterman algorithm  
Run on Mon Mar 15 07:51:39 1999 MaxPar time 776.99 Seconds  
964 996 Million cell updates/sec  
Tabular output not generated  
Title: >US-09-020-716-4  
Description: (1-5811) from US09020716 seq  
Perfect Score: 5811  
N.A. Sequence: 1 TGGGCGGTTTGGTTATGAC TATACAGAGGCGCTTGGTC 5811  
Comp: AGCGGCGAAGGCACTACTG ATAGTGCTCGGGAAGCAG

Scoring table: TABLE default  
Gap 5

Nmatch STD : Dbase 0: Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-geneseq32  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39 40:part40

Statistics: Mean 10.426; Variance 0.245; scale 1.669

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No	Score	Query Match	Length	ID	Description	Pred. No.
C	1	2242	40.7	4453	32	Human cyclin D1-human	0.00e+00
C	2	2241	40.7	4454	30	Plasmid pGEM1.8 cont	0.00e+00
C	3	2242	40.7	4510	22	Human cyclin D1-human	0.00e+00
C	4	2242	40.7	4883	8	pDEL10, plasmid DNA r	0.00e+00
C	5	2242	40.7	4883	8	pDEL10, plasmid DNA r	0.00e+00
C	6	2241	40.7	5298	4	pDEL10, plasmid DNA r	0.00e+00
C	7	2242	40.7	5298	4	pDEL10, plasmid DNA r	0.00e+00
C	8	2242	40.7	5620	9	Plasmid pVEL108 used i	0.00e+00
C	9	2242	40.7	5620	9	Plasmid pVEL108 used i	0.00e+00
C	10	2241	40.7	5555	9	Plasmid pVEL108 used i	0.00e+00
C	11	2242	40.7	7053	9	Sequence of pTE4.	0.00e+00
C	12	2240	40.6	3851	12	Plasmid pSBC-cyt/c DN	0.00e+00
C	13	2239	40.6	4045	13	Plasmid pN211 contai	0.00e+00

14	2235	40.6	4118	35	T69189	Construct pGEM-hTR(Ka	0.00e+00
15	2235	40.6	4118	35	T69188	Construct pGEM-hTR(Ka	0.00e+00
16	2235	40.6	4539	14	Q87347	Plasmid pENV1.	0.00e+00
17	2235	40.6	4621	32	T62982	Human cyclin D1-human	0.00e+00
18	2235	40.6	4621	32	T62982	Human cyclin D1-human	0.00e+00
19	2234	40.6	4234	23	V12347	Shuffled cDNA cont	0.00e+00
20	2238	40.6	8710	17	O89650	Plasmid pWPS 1.2-1.2	0.00e+00
21	2238	40.6	9020	17	T07652	pS11 FIPV EL1.	0.00e+00
22	2235	40.6	10930	20	Q81336	Plasmid pM16-1.	0.00e+00
23	2235	40.6	10930	20	Q81335	Plasmid pM16-1.	0.00e+00
24	2231	40.5	4283	37	T86610	Epidermal growth fact	0.00e+00
25	2231	40.5	4283	35	T86449	DNA encoding hGH. BGF	0.00e+00
26	2232	40.5	4713	39	V12067	Murine IAD alpha chain	0.00e+00
27	2232	40.5	4713	39	V12067	Murine IAD alpha chain	0.00e+00
28	2232	40.5	4713	39	V12067	Murine IAD alpha chain	0.00e+00
29	2232	40.5	5293	6	Q34111	Plasmid pK3.	0.00e+00
30	2230	40.5	5010	2	Q11998	Human plasminogen wit	0.00e+00
31	2234	40.5	6350	23	T35109	Plasmid pUT19-SV3.	0.00e+00
32	2234	40.5	6350	23	T35109	Plasmid pUT19-SV3.	0.00e+00
33	2234	40.5	6350	23	T35109	Plasmid pUT19-SV3.	0.00e+00
34	2226	40.4	4163	5	Q32345	Template clone Fab N9	0.00e+00
35	2226	40.4	4163	5	Q32345	Template clone Fab N9	0.00e+00
36	2225	40.4	4603	1	Q04010	Plasmid pSUN387.	0.00e+00
37	2228	40.4	4910	22	T15387	pREP76/52 hybrid aden	0.00e+00
38	2228	40.4	4950	38	V03801	Patoviral vector con	0.00e+00
39	2228	40.4	5033	12	Q74566	pB114.2 vector cont	0.00e+00
40	2228	40.4	5376	4	Q27951	pVEL149.	0.00e+00
41	2224	40.4	7635	31	T71323	Hybrid activation vec	0.00e+00
42	2224	40.4	9534	30	T62072	Vector pR124E2-emo3d	0.00e+00
43	2228	40.4	9632	38	T97898	Baculovirus transfor	0.00e+00
44	2226	40.4	19307	20	T27558	Shuttle vector pADBel	0.00e+00
45	2222	40.3	2755	5	Q30566	Vector pNN03	0.00e+00

ALIGNMENTS

RESULT 1  
ID T69890 standard: DNA: 4453 BP.

AC T69890:  
DI 19-SEP-1997 (first entry)  
DE Human cyclin D1-human CDK4 gene fusion in plasmid pK485.  
KW Cyclin D1; cyclin-dependent protein kinase 4; CDK4; cell cycle; ss.  
OS Chimeric Homo sapiens;  
CH Chimeric Homo sapiens;  
FH Key Location/Qualifiers  
FT cds 2560..4248  
FT CDS /\*tag= a  
PN WO9725345-A1.  
PD 17-JUL-1997.  
PF 03-JAN-1997: U00140.  
PR 05-JAN-1996: US-009629.  
PA (ELIL ) LILLY & CO ELI.  
PI Kovacevic S, Otto KA, Rao RN;  
DR WPI: 97-372814/34.  
DR P-PSDB: W18572.  
PT Fusion protein comprising human cyclin and cyclin dependent protein  
PT kinase - useful for research in cell cycle regulation allowing  
PT addition of both components simultaneously, giving greater control;  
PT over reaction conditions  
PS Disclosure: Page 10-15: 57pp; English.  
QC A DNA sequence (T69890) in plasmid pK485, deposited in E. coli as  
CC NPT: B-21492, includes a coding region for a fusion protein  
CC (W18572) comprising human cyclin D1 linked to human CDK4, flanked  
CC by N-terminal histidine residues, a myc epitope and a thrombin  
CC cleavage site, and by a C-terminal streptavidin binding domain  
CC to facilitate purification. The fusion protein can be expressed  
CC in host cells, pref. using a baculovirus expression system, for use  
CC in research into cell cycle regulation.  
SQ Sequence 4453 BP, 1058 A, 1172 C, 1212 G, 1011 T;  
Query Match 40.7%; Score 2242, BS 32, Length 4453;  
Best Local Similarity 95.8%, Pred. No. 0.00e+00;  
Matches 2246, Conservative 0, Mismatches 4, Indels 0, Gaps 0

113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000





















[illegible]







Query Match 40.7% Score 2241 DB 9 Length 6555.  
Best Local Similarity 100.0% P-Id No 6 00e+00.  
Matches 2241 Conservative 0 Mismatches 0 Indels 0 Gaps 0

```
Db 4315 caaacttgcgttaatactggttcataactgttccctgtgtgaataattgttatccgcgtcacaa 4374
QY 3271 CAAGCTTGGCGTAATCATGGTCATAGCTGTTCCCTGTGTGAAATGTTATCCGCTCACAA 3330
Db 4375 ttccacacacatacagccggaagcacaataagtgtaaaagctggggtgccttaataagaga 4434
QY 3331 TTCCACACACATACAGCGCGCAAAATATAAAGTGAAGAGCTTGGGTCCTTAATGAGTGA 3390
Db 4435 gtaactcacataattggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4494
QY 3391 GTTAAATACATTAATATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3450
Db 4495 gcaagctcacataataatcgcccaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 4554
QY 3451 GTCAGTGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3510
Db 4555 cttccgcttccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4614
QY 3511 GTTCCGTTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3570
Db 4615 caactcaactcaaaagcggttaatacgtttatccacaaatcagggtataacgcaggaaga 4674
QY 3571 CAAGTCACTCAAAAGCGGTAAATAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 3630
Db 4675 acatgtgaacaaagccagcaaaagccagcaaaagccagcaaaagccagcaaaagccagcaaa 4734
QY 3631 ACATGTGACCAAAAGCGGTAAATAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 3690
Db 4735 ttctccataggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4794
QY 3691 TTTTCCATAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3750
Db 4795 ggcgaacacagacagacataaagatacagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4854
QY 3751 GGCGAACACAGACAGACATAAAGATACAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3810
Db 4855 gctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4914
QY 3811 GCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3870
Db 4915 gctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4974
QY 3871 GCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3930
Db 4975 ccaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5034
QY 3931 CCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3990
Db 5035 actatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5094
QY 3991 ACTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4050
Db 5095 gtaacagattagcagcagagatattaggggggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5154
QY 4091 GTAACAGATTAGCAGCAGAGATATTAGGGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4110
Db 5155 ctaactacgctcacatagaaagcagatttgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5214
QY 4111 CTAACACGCTCACATAGAAAGCAGATTTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4170
Db 5215 cctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5274
QY 4171 CCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4230
Db 5275 gttttttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5334
QY 4231 GTTTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4290
Db 5335 tgaatcttctacaggggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5394
```

```
QY 4291 IGATCTTTTCTAGGGGTCTGAGGCTCAGTGGAGCAAAATTAACGCTTAAAGGATTTTGG 4350
Db 5395 tcatgaattatcaaaaagagatttccacctagatctcttttaaattaaaaatgaagtitta 5454
QY 4351 TCATGAGATTATCAAAAAGGATTTTCCCTAGATCTCTTTTAAATTAATAAATGAAGTTTA 4410
Db 5455 aatcaactaaagtatatatgaataaacttggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5514
QY 4411 AATCAATCTTAAATATATATGATTAACCTTGGTGTGACAGTATACCAATGCTTAAATAGTG 4470
Db 5515 aggcacattatctcagcagattgtctatttgcgttccatccatagttgtgtgtgtgtgtgtgtgt 5574
QY 4471 AGGCACCTATCTCAGCGATGCTGCTATTTCCTTCATCCATAGTGTGCTGATCTGCTGCTG 4530
Db 5575 tgaataaactacagatcagggagggttaacattctgcccacatctgcccacatctgcaatctacccg 5634
QY 4531 TGTAGATAACTAGATACGGGAGGGGTATCATCTGGCTGCTCACTGGTGAATGATACGG 4590
Db 5635 gagaccacacgtctcagcgtctccacatttaacgcaataaaacacagcagccggaagagcgcg 5694
QY 4591 GAGACCCACGCTCACCGGCTCCCAATTTATCAGTAAATAAACTAACTAACTAACTAACTAA 4650
Db 5695 agcgcaagaatggtctcgtgcacatttatccgctccatccatccatccatccatccatccatccat 5754
QY 4651 AGCGCAGAAAGTGTCTGCAACTTTATCCGCTCCATCCCACTGATTAATTTGTTGCTGGG 4710
Db 5755 aagctcagaataaagttagtctccaggttaataagtttgccgcaacggtgtgtgtgtgtgtgtgt 5814
QY 4711 AAGCTAGAGTAAGTAGTTCGCCAGTAAATAGTTTGGCAACGGTGTGTGTGTGTGTGTGTGT 4770
Db 5815 gcatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5874
QY 4771 GCATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4830
Db 5875 caagcgaggttaacatgattcccccatgttgtgcacaaagcggttagtctctctccctccctcc 5934
QY 4831 CAAGCGGAGTTATATGATTCGCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4890
Db 5935 cgatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5994
QY 4891 CGATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4950
Db 5995 ataattcttcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 6054
QY 4951 ATAATTCTCTTACTGTCTATGCCATCGGTAAATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5010
Db 6055 ccaagctcattctcaagaatagttatgagcgagcagcagcagcagcagcagcagcagcagcagcag 6114
QY 5011 CCAAGTCAATTCTGAGATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5070
Db 6115 gggataataacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 6174
QY 5071 GGGATAATAACGCGCCACATAGCAGAACTTTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5130
Db 6175 cggggcgaataactctcaagatcttaccggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 6234
QY 5131 CGGGCGGAAAACTCTCAAGGATGCTTACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5190
Db 6235 gtgaccccaactgattctcagcattcttacccttccacacagcgtttcttggtgtgtgtgtgtgtgt 6294
QY 5191 GTGACCCCACTGATCTTCAGCATCTTTTATTTTCACTAGCTGTGTGTGTGTGTGTGTGTGTGTGT 5250
Db 6295 caggaaagcgaataatgcgcgaataaaggaataagggcgacacaggaatagtgtgtgtgtgtgtgt 6354
QY 5251 CAGGAAGCGGAAATTCGCGCAAAAAAGGGAATAAGGCGACACGGAATAATTTTGAATATCTCA 5310
Db 6355 tactctctcttttcaatattattgaagcatttatcaggtttattgtctctcatgagcagat 6414
QY 5311 TACTCTCTCTTTTCAATATTATTGAAAGCATTTATCAGGCTTATTTGCTGTGTGTGTGTGTGTGTGT 5370
Db 6415 acatatttgaatgtatttgaataataaataaataaataaataaataaataaataaataaataaataa 6474
```



QY 4470 GAGGACACCAATGACGAGCAATGCTTATTTCTTTCATCATAGTGTGCTGACTCCCGCTC 4529  
 Db 6068 ggtgaataactacacatacaggaagagcttacccatctgcccacagtgctgcaataataccg 6127  
 QY 4530 GTGTACATAACTACCAIACGGGAGAGGATTAACATCTGAGGACAGTGTGCAATGATACCG 4589  
 Db 6128 ggaatccacagcttccacagcttccacatattatcagcaataaaacccacagccggaaggcc 6187  
 QY 4590 GAGACGACGCTCAACGCTAT 4649  
 Db 6188 ggaacgacgaagtggtct 6247  
 QY 4650 GAGCGAGAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4709  
 Db 6248 gaagctaaagtaagtagtgcgcacatttaattgttgcgaacgcttcttgcctactactaca 6307  
 QY 4710 GAAGCTAGAGTAAGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4769  
 Db 6308 ggcacatct 6367  
 QY 4770 GGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4829  
 Db 6368 tgaagcagagttacagatccccctgctgtgtgcacaaaagcaggttagctctctctctct 6427  
 QY 4830 TAAAGGAGATTTACATGATCCCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4889  
 Db 6428 cgcagctgttgcaggaagtagtgcgcagctgttactactctctctctctctctctctctct 6487  
 QY 4890 CGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4949  
 Db 6488 catatct 6547  
 QY 4950 CATAAATCTCTTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5009  
 Db 6548 acaagctattctgaataagtagtgcgcagcagcagctctctctctctctctctctctctct 6607  
 QY 5010 ACAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5069  
 Db 6608 ggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 6667  
 QY 5070 GCGGATAATACGCGGCAATAGCAAGCTTTTAAAGTCTCTATCATTTGGAAGAGTCTCT 5129  
 Db 6668 tccggcgcacaaactctcaagagcttaccgctgttgagatccagttcgatgtacccact 6727  
 QY 5130 TCGGGGCAAAAATCTCAAGGATCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5189  
 Db 6728 cgtgcaccccaactgatctccagcactcttctacttccacacagcttctctggttgagcaaaa 6787  
 QY 5190 CCGTGAATCAATGATCTCTTCT 5249  
 Db 6788 acaggaagcacaataagcgcacaaaaggaataaagggcagacgaaaatgttgaaatactc 6847  
 QY 5250 ACAGTAAGGCAAAATGCGCTGCAAAAAGGCAATTAAGGCGCACACCGGAAATGTTGAATCT 5309  
 Db 6848 atact 6907  
 QY 5310 ATACT 5369  
 Db 6908 tcatatttgatgtatttagaaaaataaacaataagggttcccgccagacatttcccccga 6967  
 QY 5370 TACATATTGATGATTATTAGAAAAATAAACAATAGGGTTCCCGGCACATTTCCTCCCGA 5429  
 Db 6968 aaagtccacctgacactcaaaacccatttatcatgaacattaacctataaaaatagg 7027  
 QY 5430 AAGTGCACCTGAGCTCTAAGAACCCATTATTATCATGACATTAACTATAAAAATAGG 5489  
 Db 7028 catatcacagaccccttctctc 7049  
 QY 5490 CGTATCACAGGCGCTTTCGTC 5511

RESULT 12  
 ID C67221 standard: DNA; 3854 BP.

AC Q67221:  
 DT 24-MAR-1995 (first entry)  
 DE Plasmid pSEC-cyt/c DNA encoding pre-apo-protein.  
 KW Plasmid pSEC-cyt/c; vector: halo-protein; pre-apo-protein;  
 OS protein synthesis; Escherichia coli; cytoplasm cytochrome: ss.  
 FH Synthetic.  
 FH Key Location/Qualifiers  
 FT misc\_feature 396..454  
 FT /tag= a  
 FT /label= polylinker\_site  
 FT 2948..2962  
 FT /tag= b  
 FT /label= pho\_promoter  
 FT 2996..3001  
 FT /tag= c  
 FT /label= Shine-Dalgarno\_sequence  
 FT 3009..3074  
 FT /tag= d  
 FT /note= "Alkaline phosphatase signal peptide"  
 FT 3075..3374  
 FT /tag= f  
 FT /note= "Soluble core of cytochrome-b5"  
 FT 3375..3482  
 FT /tag= g  
 FT /note= "tail of cytochrome-b5"  
 FT W09417191-A.  
 PN 04-AUG-1994.  
 PD 27-JAN-1994; G00161.  
 PF 27-JAN-1993; GB-001553.  
 PR (UYWA-) UNIV COLLEGE WALES.  
 PA Kaderbhai MA;  
 PI WPI; 94-264104/32.  
 DR P-PSDB; R57733.  
 PT Genetic precursor unit causing periplasmic translocation of  
 FT pre-apo protein - for processing the halo-protein, useful in  
 PT protein synthesis, assay of signal peptidase and identification  
 PT of its inhibitors  
 PS Claim 10; Fig. 5; 32pp; English.  
 CC This plasmid contains a DNA sequence encoding the pre-form of an apo-  
 CC protein such that the pre-apo-protein is translocated from a  
 CC cytoplasmic cell region to a periplasmic region where formation of  
 CC process apo-protein and conversion to halo-protein to occur. The  
 CC DNA encoded in this plasmid is expressed in the cytoplasm of E. coli.  
 CC especially TB-1 and M4830-1. The apo-protein is a cytoplasmic  
 CC cytochrome, especially having a soluble core domain of cytochrome-b5  
 CC of liver endoplasmic reticulum. The sequence also encodes an E.  
 CC coli alkaline phosphatase signal peptide.  
 CC Sequence 3854 BP; 1005 A; 938 C; 965 G; 946 T;  
 SQ  
 Query Match 40.6%; Score 2240; DB 12; Length 3854;  
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
 Matches 2241; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Db 445 gcaagcttgagcgaatacgttcagctgttctctctctctctctctctctctctctctctca 504  
 QY 3270 GCAAGCTTGGCGTAATCTATGCTATAGCTCTTTCTCTGTGTAATTTGTAACGCTCACA 3329  
 Db 505 attccacacacatacagcgcggaagcacaataadgttaaacctcggggtgctaatagtg 564  
 QY 3330 ATTCCACACAACATACAGCGCGGAGCAATAAGTGTAAAGCTTGGGTGCTTAATGAGTG 3389  
 Db 565 agtaactcacaataattgcttgctgctactgcccgccttccagtgagggaacacctgctg 624  
 QY 3390 AGCTAACTACATTAATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3449  
 Db 625 tgcacgtcgtatgaataac 684  
 QY 3450 TGCACGCTGCATTAATGAATCGGCTAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3509  
 Db 685 tcttcgccttct 744  
 QY 3510 TCTTCGCTTCT 3569

804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

4650  
4651  
4652  
4653  
4654  
4655  
4656  
4657  
4658  
4659  
4660  
4661  
4662  
4663  
4664  
4665  
4666  
4667  
4668  
4669  
4670  
4671  
4672  
4673  
4674  
4675  
4676  
4677  
4678  
4679  
4680  
4681  
4682  
4683  
4684  
4685  
4686  
4687  
4688  
4689  
4690  
4691  
4692  
4693  
4694  
4695  
4696  
4697  
4698  
4699  
4700  
4701  
4702  
4703  
4704  
4705  
4706  
4707  
4708  
4709  
4710  
4711  
4712  
4713  
4714  
4715  
4716  
4717  
4718  
4719  
4720  
4721  
4722  
4723  
4724  
4725  
4726  
4727  
4728  
4729  
4730  
4731  
4732  
4733  
4734  
4735  
4736  
4737  
4738  
4739  
4740  
4741  
4742  
4743  
4744  
4745  
4746  
4747  
4748  
4749  
4750  
4751  
4752  
4753  
4754  
4755  
4756  
4757  
4758  
4759  
4760  
4761  
4762  
4763  
4764  
4765  
4766  
4767  
4768  
4769  
4770  
4771  
4772  
4773  
4774  
4775  
4776  
4777  
4778  
4779  
4780  
4781  
4782  
4783  
4784  
4785  
4786  
4787  
4788  
4789  
4790  
4791  
4792  
4793  
4794  
4795  
4796  
4797  
4798  
4799  
4800  
4801  
4802  
4803  
4804  
4805  
4806  
4807  
4808  
4809  
4810  
4811  
4812  
4813  
4814  
4815  
4816  
4817  
4818  
4819  
4820  
4821  
4822  
4823  
4824  
4825  
4826  
4827  
4828  
4829  
4830  
4831  
4832  
4833  
4834  
4835  
4836  
4837  
4838  
4839  
4840  
4841  
4842  
4843  
4844  
4845  
4846  
4847  
4848  
4849  
4850  
4851  
4852  
4853  
4854  
4855  
4856  
4857  
4858  
4859  
4860  
4861  
4862  
4863  
4864  
4865  
4866  
4867  
4868  
4869  
4870  
4871  
4872  
4873  
4874  
4875  
4876  
4877  
4878  
4879  
4880  
4881  
4882  
4883  
4884  
4885  
4886  
4887  
4888  
4889  
4890  
4891  
4892  
4893  
4894  
4895  
4896  
4897  
4898  
4899  
4900  
4901  
4902  
4903  
4904  
4905  
4906  
4907  
4908  
4909  
4910  
4911  
4912  
4913  
4914  
4915  
4916  
4917  
4918  
4919  
4920  
4921  
4922  
4923  
4924  
4925  
4926  
4927  
4928  
4929  
4930  
4931  
4932  
4933  
4934  
4935  
4936  
4937  
4938  
4939  
4940  
4941  
4942  
4943  
4944  
4945  
4946  
4947  
4948  
4949  
4950  
4951  
4952  
4953  
4954  
4955  
4956  
4957  
4958  
4959  
4960  
4961  
4962  
4963  
4964  
4965  
4966  
4967  
4968  
4969  
4970  
4971  
4972  
4973  
4974  
4975  
4976  
4977  
4978  
4979  
4980  
4981  
4982  
4983  
4984  
4985  
4986  
4987  
4988  
4989  
4990  
4991  
4992  
4993  
4994  
4995  
4996  
4997  
4998  
4999  
5000

RESULT 13

ID Q70942 standard; DNA: 4045 BP.

AC Q70942;

DT 22-MAR-1995 (first entry)

DE Plasmid pND211 contained synthetic human trophoblastin (SHBG) gene.

KW Trophoblastin; pharmaceutical; surgical dressing; plasmid pND211; ds.

OS Synthetic.

PN W09414958-A.

PU 07-JUL-1994.

PF 16-DEC-1993; A00655.

PR 22-DEC-1992; A006520.

PR 28-JUN-1993; A009661.

PA (UNSY ) UNIV SYDNEY.

PI	Martin SL, Weiss AS.	PI	Martin SL, Weiss AS.
DP	Wf: 94-253633/32	DP	Synthetic polynucleotide(s) - encode recombinant tropoelastins
PT	and variants	PT	and variants
PS	Disclosure: Page 48; 77pp; English.	PS	Disclosure: Page 48; 77pp; English.
CC	This nuc-tion vector has the human synthetic tropoelastin gene	CC	This nuc-tion vector has the human synthetic tropoelastin gene
CC	inserted into an EcoI site and uses phage lambda promoter-driven	CC	inserted into an EcoI site and uses phage lambda promoter-driven
CC	expression to permit initiation from the start codon of the	CC	expression to permit initiation from the start codon of the
CC	synthetic gene. The expressed product is susceptible to	CC	synthetic gene. The expressed product is susceptible to
CC	hydrolytic breakdown of the crosslinks. Such material may be	CC	hydrolytic breakdown of the crosslinks. Such material may be
CC	used in e.g. surgical applications, where the gradual loss of	CC	used in e.g. surgical applications, where the gradual loss of
CC	material over time is intended.	CC	material over time is intended.
SC	Sequence 4045 bp; 1045 A; 994 C; 956 G; 1049 T;	SC	Sequence 4045 bp; 1045 A; 994 C; 956 G; 1049 T;
Query: March 40 58; Score 2239; DB 12; Length 4045;			
Best Local Similarity 100.0%; Pired No 0 000-00;			
Matches 2240; Conservative 0; Mismatches 1; Indels 0; Gaps 0			
Db	399 gacgaagggcctcgtgatacgcctattttataggttaatgctcatgataaatggttt	458	gacgaagggcctcgtgatacgcctattttataggttaatgctcatgataaatggttt
Cp	5511 GACGAAAGGGCTCGTGATACGGCTATTTTATAGGTTAAATGTCATGATATATGGTTT	5452	GACGAAAGGGCTCGTGATACGGCTATTTTATAGGTTAAATGTCATGATATATGGTTT
Db	459 cttaaacgtcaggtggcacttttcgggaaatgtgcgggaacccctattgtttatttt	518	cttaaacgtcaggtggcacttttcgggaaatgtgcgggaacccctattgtttatttt
Cp	5451 CTATGACGTCTAGTGTGTGATTTTGGGGAAATGTGGGGGAACCCCTATTGTATTATTT	5392	CTATGACGTCTAGTGTGTGATTTTGGGGAAATGTGGGGGAACCCCTATTGTATTATTT
Db	519 tctaaatacatccaatgatgtatccctcatgagacaataaccctgataaatgctcaat	578	tctaaatacatccaatgatgtatccctcatgagacaataaccctgataaatgctcaat
Cp	5391 TCTAATACATTCAANTATGTATCCCTCTATGAGACATATACCCCTGATATAATGCTTCAAT	5332	TCTAATACATTCAANTATGTATCCCTCTATGAGACATATACCCCTGATATAATGCTTCAAT
Db	579 atatgtgaaaaggagagatgagatttcaacatttcctgtgcgcctattccctttt	638	atatgtgaaaaggagagatgagatttcaacatttcctgtgcgcctattccctttt
Cp	5331 AATATTGAAAAAGGAGAGATGAGTATTCACATTTCCGTGTGCGCTTATTCCCTTTT	5272	AATATTGAAAAAGGAGAGATGAGTATTCACATTTCCGTGTGCGCTTATTCCCTTTT
Db	639 ttgcgcatttgccttctgttttgcaccagaacacctggtgaaagttaaagatg	698	ttgcgcatttgccttctgttttgcaccagaacacctggtgaaagttaaagatg
Cp	5271 TTGTCAGATTTTCCCTTCTCTGTTTITGTCACATAGAAACCTGGTGAAGTAAAGATG	5212	TTGTCAGATTTTCCCTTCTCTGTTTITGTCACATAGAAACCTGGTGAAGTAAAGATG
Db	699 ctgaagatcagttgggtgcagagtggtgtacatggaactggatctcaacagcggttaaga	758	ctgaagatcagttgggtgcagagtggtgtacatggaactggatctcaacagcggttaaga
Cp	5211 CTGAACATCATTTGGGTGTGACGAGTGGGTTCATCTGAATCGATCTCAACAGCGGTGAAGA	5152	CTGAACATCATTTGGGTGTGACGAGTGGGTTCATCTGAATCGATCTCAACAGCGGTGAAGA
Db	759 tccctgagagttttcggcccggaagaacgttttccaatgatgagcacttttaaagttctgc	818	tccctgagagttttcggcccggaagaacgttttccaatgatgagcacttttaaagttctgc
Cp	5151 TCTTTGAGAGTTTTGGCCGTGAAGAACCTTTTCCAAATGATGAGCACCTTTAAAAGTTTCG	5092	TCTTTGAGAGTTTTGGCCGTGAAGAACCTTTTCCAAATGATGAGCACCTTTAAAAGTTTCG
Db	819 tatctgcccgtattatccctattgacgcgggcaagcaactcggtgcgcgatac	878	tatctgcccgtattatccctattgacgcgggcaagcaactcggtgcgcgatac
Cp	5091 TATGTGGCGCGGTATTATCCCGTATTGACCGCGGCACAGACGACTCGTGCCTGGCATAC	5032	TATGTGGCGCGGTATTATCCCGTATTGACCGCGGCACAGACGACTCGTGCCTGGCATAC
Db	879 actattctcagaatgacttggttgagtactcaccagtcacagaaagcatctttacggatg	938	actattctcagaatgacttggttgagtactcaccagtcacagaaagcatctttacggatg
Cp	5031 ACTATTCTCAGAAATGACTTGGTTGAGTATCATCACCAGTCACAGAAAAAGCATCTTACGGATG	4972	ACTATTCTCAGAAATGACTTGGTTGAGTATCATCACCAGTCACAGAAAAAGCATCTTACGGATG
Db	939 gnatgacgtgaagaaattatgcatgtgtccataaccatgaatgataaacactgcggcca	998	gnatgacgtgaagaaattatgcatgtgtccataaccatgaatgataaacactgcggcca
Cp	4971 GATATGATATATGAAATATATGATGTTGATATACCATGAGTGAATCAACCTCGCGCCA	4912	GATATGATATATGAAATATATGATGTTGATATACCATGAGTGAATCAACCTCGCGCCA
Db	999 acttactctgacaacgatcgaggaacggaagagctaacgcgttttttgcacaacatgg	1058	acttactctgacaacgatcgaggaacggaagagctaacgcgttttttgcacaacatgg
Cp	4911 ACTTACTTCTGACAGCATGGAGACACCAAGCAAGCACTTAACCGCTTTTTTGACAAACATGG	4852	ACTTACTTCTGACAGCATGGAGACACCAAGCAAGCACTTAACCGCTTTTTTGACAAACATGG
Db	1059 ggaatcatgaactcgtctgatatcttgggaacggagactgaatgaagccataaccacaag	1118	ggaatcatgaactcgtctgatatcttgggaacggagactgaatgaagccataaccacaag
Cp	4851 GGATATATTAATCTGCTTTGATGCTTTGGGGAACCGGAGCTGAATGAAAGCCATACCAACG	4792	GGATATATTAATCTGCTTTGATGCTTTGGGGAACCGGAGCTGAATGAAAGCCATACCAACG
Db	1119 acgagcgtgacaccacgatgcctgttagcaatggcaacaacgttgcgcaactattaactg	1178	acgagcgtgacaccacgatgcctgttagcaatggcaacaacgttgcgcaactattaactg
Cp	4791 ACGAGGCTGACACACGATGTCTCTGTAGCAATGGCAACACCGTTGCGCAAACTATTAACTG	4732	ACGAGGCTGACACACGATGTCTCTGTAGCAATGGCAACACCGTTGCGCAAACTATTAACTG
Db	1179 gcaactacttactctagcttcccgcgcaacaattaatagactgtagtggcgcgataaag	1238	gcaactacttactctagcttcccgcgcaacaattaatagactgtagtggcgcgataaag
Cp	4731 GGAAGTACTTACTTAGCTTCCGGCAACAATTAATAGACTGGATGGAGCGGATAAAG	4672	GGAAGTACTTACTTAGCTTCCGGCAACAATTAATAGACTGGATGGAGCGGATAAAG

[illegible]



Db 2086 atgaattatcaaaagaagatcttcaactagatccttttaataattaaaaatgaagttttaa 2145  
QY 4353 ATGACATATCAAAAAGGATCTTCACCTAGATGCTTTTAAATTAANAATGAAGTTTAAA 4412  
Db 2146 tcaatcaaatatataatgaataacttggtgtgacagttaccatgtcttaatacagtgag 2205  
QY 4413 TCAATCAAAATATATAGTAAGTAAGTTGGTCTGACAGTTACCAATGCTTAATCAGTGAG 4472  
Db 2206 gacactatctcagcagatctctctattctcattctcattccatagttgactccctcgctg 2265  
QY 4473 GCACCTATCTAGCCATCTGTCTCTATTCTGTTCTATCTCATAGTTGCTGACTCCCTCGTGTG 4532  
Db 2266 tagataactcagatcaggaagggctaccatctctggccccagtgctgcaatgataccgcga 2325  
QY 4533 TAGATAACTATGATACGAGAGGGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4592  
Db 2326 gaccacactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2385  
QY 4593 GACCCACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTG 4652  
Db 2386 cgcgaatggtcctgcgaacttaccgctctccatccagctctattatgtgtgctggagaa 2445  
QY 4653 GCGCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4712  
Db 2446 gctagatgaagttagtgcgcagtttaagtgttgccgaacagttgttgcattgctacagggc 2505  
QY 4713 GCTAGACTAAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4772  
Db 2506 atcgtgactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2565  
QY 4773 ATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4832  
Db 2566 agcagagttacatgaccccatgctgtgcgaagagcaggttagctccttcgctccctcgc 2625  
QY 4833 AGCGAATTAAT 4892  
Db 2626 atcgttctcagagtaadttagccagcagtttgcacatcagttggttatggcagcagctgcat 2685  
QY 4893 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4952  
Db 2686 aattctctactgctacgctcagcagcagcagcagcagcagcagcagcagcagcagcagc 2745  
QY 4953 AATTCCTAT 5012  
Db 2746 aactctctctgagagtagtgcagcagcagcagcagcagcagcagcagcagcagcagc 2805  
QY 5013 AAGTCATTCGSAATAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5072  
Db 2806 gataatccgcgcccacatagcagaactttaaaagtgctcatttggaacagttcttctcg 2865  
QY 5073 GATAATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5132  
Db 2866 gacgcaaaactcagagatctaccgctgttgagatcacttgcatttgatataaccacactgt 2925  
QY 5133 GGGCGAANAACCTCAGGATCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5192  
Db 2926 gacgcaaaactcagagatctaccgctgttgagatcacttgcatttgatataaccacact 2985  
QY 5193 GCACCCAAATGATCTTCAGCATCTTTTACTTTTCACTAGGCTTCTGCTGCTGCTGCTG 5252  
Db 2986 ggaatgaaatgagcagcaaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaa 3045  
QY 5253 GAAGGCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5312  
Db 3046 gctctcctcttcaatattatgaagcatttatcagaggttatctctcctcctcagcagaa 3105  
QY 5313 CTCTCTCTCTTTTCAATATATATGACGATTTATATGAGGCTTATATGCTCTATGAGG 5372  
Db 3106 atattgagatgatttagaaaaataaacaataggggttcgcgcgcacatttccccgcaaaa 3165  
QY 5373 ATATTGAAATGATTATAGAAAAAIAAACAATAAGGCTTCTGCTGCTGCTGCTGCTGCTG 5432

Db 3166 gtccacactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3225  
QY 5433 GTCCACACTGAGGCTGTAGAAAGCTATTAATATGACATTAACCTAIAAAAAAIAAGGCT 5492  
Db 3226 atcacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3244  
QY 5493 ATCAGGAGGCGCTTTCGTC 5511  
RESULT 15  
ID 169188 standard: DNA; 4118 BP.  
AC 169188;  
DT 26-FEB-1998 (first entry)  
DE Construct pGEM-HTR containing RNA component of human telomerase.  
KW Human telomerase; quantification; tumour cell; pGEM-HTR;  
KW detection; micrometastasis; diagnosis; lymphoblastoma; leukaemia;  
KW teratocarcinoma; melanoma; carcinoma; cancer; tumour; neuroblastoma;  
KW rhabdomyosarcoma; leiomyosarcoma; lymphoma; RNA component; ss.  
OS Homo sapiens.  
CS Synthetic.  
FH Key Location/Qualifiers  
FT misc\_feature 12..975 FT  
FT /tag= a  
FT /note= \*cDNA from RNA component of human telomerase\*  
FN WO9718322-A2.  
PD 22-MAY-1997.  
PF 14-NOV-1996; D02183.  
PR 15-NOV-1995; DE-042795.  
PA (DAHM/) DAHM M.W.  
PI Dahm MW.  
DR WPI; 97-289298/26.  
FT Quantifying tumour cells in body fluid - by measuring RNA component.  
PT of telomerase after amplification, especially useful for early  
PT diagnosis of metastasis  
PS Example 5; Fig 5; 46pp; German.  
CC The present sequence is the construct pGEM-HTR, which comprises the  
CC transcription vector pGEM-13zf(-) and the cDNA from the RNA  
CC component of human telomerase. The construct was used in the  
CC development of a novel method for quantifying tumour cells in a  
CC body fluid. The method comprises specific amplification of the RNA  
CC component of telomerase, and measuring the amount of amplified  
CC nucleic acid. At least 1, preferably all 3 (769173-75) standard  
CC nucleic acids are present (at different concentrations) and are  
CC co-amplified with telomerase DNA. The amplification products are  
CC detected directly or via a label or by hybridisation with a  
CC labelled oligonucleotide (769184-87, 1 for each standard and 1 for  
CC telomerase), and the amount of telomerase product compared with  
CC that from the standards.  
CC The method can be used to detect tumour cells, specifically  
CC micrometastases, in a body fluid, particularly for the early  
CC diagnosis of metastatic spread and for monitoring tumour therapy.  
CC Typical tumour cells that can be quantified are micrometastases,  
CC T cell lymphoblastoma or leukaemia (chronic myelogenous, and  
CC chronic or acute lymphatic leukaemia), teratocarcinoma, melanoma,  
CC carcinoma of lung, liver or prostate, cancer of colon or breast,  
CC kidney, brain or adrenal tumours, neuroblastoma, rhabdomyosarcoma,  
CC leiomyosarcoma and/or lymphoma.  
CC Sequence 4118 BP, 531 A, 1032 C, 1094 G, 1001 T;  
SU Query Match 40.88; Score 2235, DB 35; Length 4118  
Best Local Similarity 99.94; Pred. No. 0.00e+00;  
Matches 2237; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Db 1006 agcttgccgtaacatcaggttcagctgcttccctgtgtaaatgttatccgctcacaatt 1065  
QY 3273 AGCTTGGCGTAATCATGCTGCTATAGCTGTGTTCTGTGTGAAATGTTATGCGCTCAAA 1132  
Db 1066 ccaac 1125  
QY 3333 CCACAAACATACAGGCGGAGGATATAAGCTGTAAGCTTGGGCTTCTCTAATAGTAGGCT 1192  
Db 1126 taactcacttaattcaggttgctcagcagcttccagtcaggagaaactcctcctcagc 1185







\*\*\*\*\*

W P S R E H

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_on a - s.a. database search, using Smith-Waterman algorithm  
Run on: Mon Mar 15 04:49:36 1999. MasPar time 7003.29 Seconds  
Tabular output not generated  
1409.192 Million cell updates/sec

Title: >US-09-020-716-4  
Description: (1,511) ftm US09020716.seq  
Perfect Score: 5511  
N.A. Sequence: 1 TCGCGCTTTCGTGTATGAC..... TATCAGGAGGCGCTTTCGT 5511  
Comp: AGCGCGCAAGCACTACTG ATAGTGTCTCGGCAAGCAG

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 227026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 08

Listing first 45 summaries

Database: emb1-est56

Database: genbank-est109

5\_gb-est1 6:gb-est10 7:gb-est11 8:gb-est12 9:gb-est13  
10:gb-est14 11:gb-est15 12:gb-est16 13:gb-est17  
14:gb-est18 15:gb-est19 16:gb-est20 17:gb-est21  
18:gb-est22 19:gb-est23 20:gb-est24 21:gb-est25 22:gb-est26  
23:gb-est27 24:gb-est28 25:gb-est29 26:gb-est30 27:gb-est31  
28:gb-est32 29:gb-est33

Statistics: Mean 12.756; Variance 2.337; scale 5.464

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	564	12.0	711	26	AG004608 Homo sapiens genomic D	0.00e+00
2	564	12.0	711	27	AG010947 Homo sapiens genomic D	0.00e+00
3	557	11.9	721	27	AG011001 Homo sapiens genomic D	0.00e+00
4	557	11.9	721	25	AG004552 Homo sapiens genomic D	0.00e+00
5	530	11.4	698	26	AG003787 Homo sapiens genomic D	0.00e+00
6	530	11.4	698	27	AG009976 Homo sapiens genomic D	0.00e+00
7	531	11.4	718	26	AG004363 Homo sapiens genomic D	0.00e+00
8	531	11.4	718	27	AG010489 Homo sapiens genomic D	0.00e+00
9	524	11.3	696	27	AG009765 Homo sapiens genomic D	0.00e+00
10	524	11.3	696	26	AG003576 Homo sapiens genomic D	0.00e+00
11	520	11.2	720	25	AG006062 Homo sapiens genomic D	0.00e+00
12	520	11.2	720	28	AG013958 Homo sapiens genomic D	0.00e+00
13	519	11.2	747	25	AG007052 Homo sapiens genomic D	0.00e+00

C	14	604	11.0	692	27	AG040352	CIT-HSP-2327K01.TF CIT	0.00e+00
	15	603	10.9	695	26	AG002183	Homo sapiens genomic D	0.00e+00
	16	594	10.8	640	27	AG074298	CIT-HSP-2322L21.TF CIT	0.00e+00
	17	593	10.8	703	26	AG001761	Homo sapiens genomic D	0.00e+00
	18	594	10.8	748	17	AI124281	1.53 Drosophila 8-12 h	0.00e+00
	19	589	10.7	452	28	AQ108423	CIT-HSP-2379M3.TF CIT	0.00e+00
	20	590	10.7	690	27	AG009464	Homo sapiens genomic D	0.00e+00
	21	584	10.6	666	27	AG079096	CIT-HSP-2356E34.TF CIT	0.00e+00
	22	579	10.5	594	26	C17B6	Ciona intestinalis gen	0.00e+00
	23	577	10.5	682	28	AG014394	Homo sapiens genomic D	0.00e+00
	24	568	10.3	542	28	AQ112212	CIT-HSP-2372L22.TF CIT	0.00e+00
	25	558	10.1	615	28	AG112424	CIT-HSP-2376F2.TF CIT	0.00e+00
	26	551	10.0	604	26	AG002706	Homo sapiens genomic D	0.00e+00
	27	551	10.0	634	27	AG038010	CIT-HSP-2326F2.TF CIT	0.00e+00
	28	550	10.0	633	27	AG077534	CIT-HSP-2345K01.TF CIT	0.00e+00
	29	545	9.9	633	27	AG040787	CIT-HSP-2336K22.TF CIT	0.00e+00
	30	546	9.9	661	27	AG057693	CIT-HSP-2345M2.TF CIT	0.00e+00
	31	541	9.8	637	27	AQ077350	CIT-HSP-2365E1.TF CIT	0.00e+00
	32	537	9.7	626	28	AQ111342	CIT-HSP-2371M2.TF CIT	0.00e+00
	33	535	9.7	682	27	AQ074693	CIT-HSP-2301L23.TF CIT	0.00e+00
	34	526	9.5	540	27	AG074307	CIT-HSP-2382N1.TF CIT	0.00e+00
	35	526	9.5	594	26	AG007826	Homo sapiens genomic D	0.00e+00
	36	526	9.5	600	28	AQ111431	CIT-HSP-2373F2.TF CIT	0.00e+00
	37	523	9.5	629	27	AG011367	Homo sapiens genomic D	0.00e+00
	38	526	9.5	743	26	AG002826	Homo sapiens genomic D	0.00e+00
	39	517	9.4	623	27	AQ081172	CIT-HSP-2357C21.TF CIT	0.00e+00
C	40	518	9.4	647	27	AG008943	Homo sapiens genomic D	0.00e+00
	41	514	9.3	612	27	AG080596	CIT-HSP-2382D1.TF CIT	0.00e+00
	42	510	9.3	664	27	AG012931	Homo sapiens genomic D	0.00e+00
	43	510	9.3	664	26	AG005286	Homo sapiens genomic D	0.00e+00
C	44	509	9.2	705	27	AG009803	Homo sapiens genomic D	0.00e+00
C	45	509	9.2	705	26	AG003614	Homo sapiens genomic D	0.00e+00

ALIGNMENTS

RESULT LOCUS	1	AG004608	711 bp	DNA	GSS	29-JAN-1998
DEFINITION		Homo sapiens genomic DNA, 21q region, clone: S594B38, genomic survey sequence.				
ACCESSION		AG004608				
NID		Q2822058				
KEYWORDS		GSS.				
SOURCE		Homo sapiens				
ORGANISM		Homo sapiens DNA, clone: S594B38.				
REFERENCE		1 (bases 1 to 711)				
AUTHORS		Hattori,M., Ishii,K., Shiba,T and Sakaki,Y.				
TITLE		Homo sapiens genomic DNA, chromosome 21q				
JOURNAL		Published Only in DataBase (1998) In press				
REFERENCE		2 (bases 1 to 711)				
AUTHORS		Hattori,M., Ishii,K., Shiba,T and Sakaki,Y.				
TITLE		Direct Submission				
JOURNAL		Submitted (29-JAN-1998) to the DDBJ/EMBL/GenBank databases.				
		Masahira Hattori, Kitasato University, Department of Science, 357				
		Seiichi Hattori, Kitasato University, Department of Science, 357				
		(E-mail: hattori@ipc.kit.ac.jp, Tel: 0427-78-9732,				
		Fax: 0427-78-9561)				
FEATURES		Location/Qualifiers				
source		1. /11				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
		/chromosome="21"				
		/clone="S594B38"				
		/map="21q"				
BASE COUNT		169 a 188 c 172 g 192 t				
ORIGIN						

Query Match 12.0% Score 564, DB 26; Length 711;  
Best Local Similarity 99.6%, Pred. No. 0.00e+00;  
Matches 682, Conservative 0, Mismatches 0, Indels 3, Gaps 3;



```

1 (bases 1 to 721)
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
Homo sapiens genomic DNA, Chromosome 21q
Published Only in Database (1998) In press
2 (bases 1 to 721)
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
Direct Submission
Submitted (29-JUL-1998) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@hgc.ims.u-tokyo.ac.jp; tel:0427-78-9732,
Fax:0427-78-9561)
FEATURES             Location/Qualifiers
     source            1..721
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /clone="879G12X91"
                     /map="21q"
     base_count        195 a 163 c 175 g 179 t 9 others
     origin            1..721

Query Match      11.9%  Score 657;  DB 27;  Length 721;
Best Local Similarity 98.7%  Pred. No. 0.00e+00;
Matches 683;  Conservative 0;  Mismatches 5;  Indels 4;  Gaps 4;

Db      32  CATCTTACGATGGCATGATGACGATGAACAGATATATGTCAGTCGTCGCCATAACCATGAGTGAT 91
Cp      4984  CATCTTACGATGGCATGATGACGATGAACAGATATATGTCAGTCGTCGCCATAACCATGAGTGAT 4925

Db      92  AATACATGCGGCTAACTTACTTCTGATCAACAGATGGAGGACCGAAGGAGCTAACCGGTTTT 151
Cp      4924  AACACTGCGGCCAACTTACTTCTGCAACAGCATCGAGGACCGAAGGAGCTAACCGGTTTT 4865

Db      152  TTGCACACATGSGGGATCATGTAACCTGCGCTTGATCGTTGGGAACCGGAGCTGAATGAA 211
Cp      4864  TTGCACAACTATGSGGGATCATGTAATTCGCTTATGATCGCTTGGGAACCGGAGCTGAATGAA 4805

Db      212  GGCATCAATCAACGATCAGGCTGATACATCAATGAGTGTATGATGACAACTGGCTGGCC 271
Cp      4804  GCATACCAAAAGATGAGGTGATACATCAATGAGTGTATGATGACAACTGGCTGGCC 4745

Db      272  AAACATTAATCTGGGAATCTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 331
Cp      4744  AAACATTAATCTGGCAACTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 4625

Db      332  GAGCGGATAAAGTTGGAGACCACTTCGCGCTCGGCGCTCGGCGCTCGGCGCTCGGCGCTT 391
Cp      4684  GAGCGGATAAAGTTGCAGGACCACCTTCGCGCTCGGCGCTCGGCGCTCGGCGCTCGGCGCTT 4625

Db      392  GCTGATAAATCTGAGGCGGTGAGCGGTGGTCTCGGGGTATCATTCAGCACTGGGGCCA 451
Cp      4624  GCTGATAAATCTGAGCGCGGTGAGCGGTGGTCTCGGGGTATCATTCAGCACTGGGGCCA 4565

Db      452  GATGGTAAGGCTTCTGATATGCTAGTATATCTACACACGCGGGAGTCAGCACTATGGAT 511
Cp      4564  GATGGTAAGGCTTCTGATATGCTAGTATATCTACACACGCGGGAGTCAGCACTATGGAT 4505

Db      512  GAACGAATACAGAGATCGGTACAGATAGGTCGCTCAGCTGATTAAGC-TTGGTAACCTGCA 570
Cp      4504  GAATGAATACAGAGATCGGTACAGATAGGTCGCTCAGCTGATTAAGC-TTGGTAACCTGCA 4445

Db      571  GACCAAGTTTACTC-TATATACTTTAGATTGATTTAAACCTCATTTTTTAATTTAAAGG 629
Cp      4444  GACCAAGTTTACTCATAATATCTTTAGATTGATTTAAACCTCATTTTTTAATTTAAAGG 4385

Db      630  ATCTAGGTGAAGATGCTTTT-GATAATCAGTACGCAAA-TCCCTTAACCTGAGTTTCG 687
Cp      4384  ATCTAGGTGAAGATGCTTTTGTATATCTCATGACCAAAATCCCTTAACCTGAGTTTCG 4325

Db      688  TTCCACTAGCGTCAGACCCGCTAGAAAAAT 719
Cp      4324  TTCCACTAGCGTCAGACCCGCTAGAAAAAT 4293

```





Yoshiko Hattori, Graduate University, Department of Science, 1st  
Sequences Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan  
(E-mail: hattori@ipc.ims.u-tokyo.ac.jp, Tel: 042-754-4132,  
Fax: 04-279-6561)

FEATURES	LOCATION/Qualifiers
SOURCE	1. 7718
	/organism "Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="21"
	/clone="178X25"
	/map="21q"
BASE COUNT	176 a 191 c 187 g 168 t 6 others
ORIGIN	
Query Match	11.4% Score 631; IR 27; Length 718;
Best Local Similarity	97.0% Pred. No. 0.000+00;
Matches 640; Conservative 9;	Mismatches 15;
	Indels 6 Gaps 6;

18	21	4600	4601	4602	4603	4604	4605	4606	4607	4608	4609	4610	4611	4612	4613	4614	4615	4616	4617	4618	4619	4620	4621	4622	4623	4624	4625	4626	4627	4628	4629	4630	4631	4632	4633	4634	4635	4636	4637	4638	4639	4640	4641	4642	4643	4644	4645	4646	4647	4648	4649	4650	4651	4652	4653	4654	4655	4656	4657	4658	4659	4660	4661	4662	4663	4664	4665	4666	4667	4668	4669	4670	4671	4672	4673	4674	4675	4676	4677	4678	4679	4680	4681	4682	4683	4684	4685	4686	4687	4688	4689	4690	4691	4692	4693	4694	4695	4696	4697	4698	4699	4700	4701	4702	4703	4704	4705	4706	4707	4708	4709	4710	4711	4712	4713	4714	4715	4716	4717	4718	4719	4720	4721	4722	4723	4724	4725	4726	4727	4728	4729	4730	4731	4732	4733	4734	4735	4736	4737	4738	4739	4740	4741	4742	4743	4744	4745	4746	4747	4748	4749	4750	4751	4752	4753	4754	4755	4756	4757	4758	4759	4760	4761	4762	4763	4764	4765	4766	4767	4768	4769	4770	4771	4772	4773	4774	4775	4776	4777	4778	4779	4780	4781	4782	4783	4784	4785	4786	4787	4788	4789	4790	4791	4792	4793	4794	4795	4796	4797	4798	4799	4800	4801	4802	4803	4804	4805	4806	4807	4808	4809	4810	4811	4812	4813	4814	4815	4816	4817	4818	4819	4820	4821	4822	4823	4824	4825	4826	4827	4828	4829	4830	4831	4832	4833	4834	4835	4836	4837	4838	4839	4840	4841	4842	4843	4844	4845	4846	4847	4848	4849	4850	4851	4852	4853	4854	4855	4856	4857	4858	4859	4860	4861	4862	4863	4864	4865	4866	4867	4868	4869	4870	4871	4872	4873	4874	4875	4876	4877	4878	4879	4880	4881	4882	4883	4884	4885	4886	4887	4888	4889	4890	4891	4892	4893	4894	4895	4896	4897	4898	4899	4900	4901	4902	4903	4904	4905	4906	4907	4908	4909	4910	4911	4912	4913	4914	4915	4916	4917	4918	4919	4920	4921	4922	4923	4924	4925	4926	4927	4928	4929	4930	4931	4932	4933	4934	4935	4936	4937	4938	4939	4940	4941	4942	4943	4944	4945	4946	4947	4948	4949	4950	4951	4952	4953	4954	4955	4956	4957	4958	4959	4960	4961	4962	4963	4964	4965	4966	4967	4968	4969	4970	4971	4972	4973	4974	4975	4976	4977	4978	4979	4980	4981	4982	4983	4984	4985	4986	4987	4988	4989	4990	4991	4992	4993	4994	4995	4996	4997	4998	4999	5000	5001	5002	5003	5004	5005	5006	5007	5008	5009	5010	5011	5012	5013	5014	5015	5016	5017	5018	5019	5020	5021	5022	5023	5024	5025	5026	5027	5028	5029	5030	5031	5032	5033	5034	5035	5036	5037	5038	5039	5040	5041	5042	5043	5044	5045	5046	5047	5048	5049	5050	5051	505
----	----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----

PERSON: AG009765 696 bp DNA GSS 06-JUL-1998  
 ORGANISM: Homo sapiens chromosomal DNA, 240,000 bp genomic  
 survey sequence.  
 ACCESSION: AG009765  
 NID: 33284751

KEYWORD SUBJECT ORGANISM

GSS.  
Homo sapiens DNA, cDNA; pnc4; pN15.  
Homo sapiens  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

AUTHORS  
Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.  
TITLE  
Homo sapiens genomic DNA, chromosome 21p  
JOURNAL  
Published only in Databank (1998) In press  
REFERENCE  
2 (bases 1 to 996)  
AUTHORS  
Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.  
JOURNAL  
Direct Submission  
TITLE  
Submitted (06-JUN-1998) to the GenBank database.  
Masahiro Hattori, Kitasato University, Department of Biogenetics,  
Sequencing Laboratory, Kitasato 1-15-1, Saitama 426, Japan.  
(E-mail: hattori@pc.ims.u-tokyo.ac.jp, tel: 0427 78 9797,  
Fax: 0427 78 9561)

[illegible]



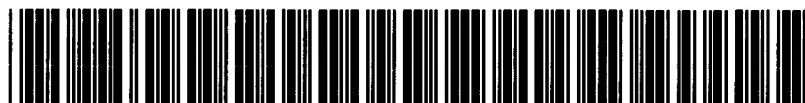








US 090207161CP1



Creation date: 28-08-2003  
Indexing Officer: FNIGATU - FITSUM NIGATU  
Team: OIPEBackFileIndexing  
Dossier: 09020716

Legal Date: 18-05-1999

No.	Doccode	Number of pages
1	NPL	8
2	NPL	14
3	NPL	4
4	NPL	10

Total number of pages: 36

Remarks:

Order of re-scan issued on .....

